View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Sissa Digital Library

SCUOLA INTERNAZIONALE SUPERIORE di STUDI AVANZATI
International School for Advanced Studies
SISSA

Optimal transportation in geodesic spaces

Ph.D. Thesis

Supervisor
Prof. Stefano Bianchini

Candidate
Fabio Cavalletti

ACADEMIC YEAR 2010/2011

Il presente lavoro costituisce la tesi presentata da Fabio Cavalletti, sotto la direzione di ricerca del prof. Stefano Bianchini, al fine di ottenere l'attestato di ricerca post-universitaria di *Doctor Philosophiae in Analisi Matematica* presso la SISSA di Trieste. Ai sensi dell'art. 18, comma 3, dello Statuto della SISSA pubblicato sulla G.U. n. 62 del 15.03.2001, il predetto attestato è equipollente al titolo di *Dottore di Ricerca in Matematica*.

Trieste, anno accademico 2010/2011.

Contents

Introduction

The optimal transportation problem was firstly formulated and studied by Monge in 1781 and it has generated in the last years an important branch of mathematics.

The problem originally studied by Gaspard Monge was the following: in $\mathbb{R}^3$ assume that we are given a pile of sand and a hole that we have to fill up completely with the sand. Clearly the pile and the hole must have the same volume and different ways of moving the sand will give different costs of the operation. Monge wanted to minimize the cost of this operation.

Nowadays the Monge transportation problem can be stated in the following general form: given two probability measures μ and ν, defined on the measurable spaces X and Y, find a measurable map $T : X \to Y$ with

$$T_\sharp \mu = \nu, \qquad \text{i.e.} \quad \nu(A) = \mu(T^{-1}(A)) \quad \forall A \subset Y \text{ measurable}$$

in such a way that T minimizes the transportation cost, that is

$$\int_X c(x, T(x))\mu(dx) = \min_{\hat{T}_\sharp \mu = \nu} \int_X c(x, \hat{T}(x))\mu(dx),$$

where $c : X \times Y \to \mathbb{R}^+$ is some given cost function and the minimum is taken over all measurable map $\hat{T} : X \to Y$ such that $\hat{T}_\sharp \mu = \nu$. When the transport condition $T_\sharp \mu = \nu$ is satisfied, we say that T is a transport map, and if T minimizes also the cost we call it an optimal transport map.

An important part of this thesis is devoted to the study of the existence of optimal transport map. In particular we will present some results regarding the existence of optimal transportation map in the abstract framework of metric spaces with geodesic distance cost under very mild non-degeneracy conditions on the first marginal measure μ.

During the last decades, techniques introduced for studying optimal transportation problems have found many application to other branches of mathematics such as PDE's or metric geometry. We will focus on the synthetic formulation of lower bounds on Ricci curvature in terms of optimal transportation inequality. In 2006 Sturm and independently Lott and Villani presented a concept of lower "Ricci" curvature bound linked with a generalized upper bound on the dimension in the setting of abstract metric measure spaces (M, d, m), the so-called *curvature-dimension condition* $\mathsf{CD}(K, N)$. The parameter K and N play the role of curvature and dimension bound. The definition is based on convexity properties of the entropy as function on the Wasserstein space $\mathcal{P}_2(M, d)$ of probability measure on the metric space (M, d).

It is still not known whether this notion satisfies a globalization property, i.e. assume that every point of the space has a neighborhood satisfying $\mathsf{CD}(K, N)$ with the curvature bound independent on the point, then the whole space (M, d, m) satisfies $\mathsf{CD}(K, N)$. In the last part of this thesis we will present a partial globalization result for showing that if a metric measure space satisfies the local version of $\mathsf{CD}(K, N)$, then it satisfies a curvature and dimension bound called measure contraction property, $\mathsf{MCP}(K, N)$, slightly weaker, in our setting, of $\mathsf{CD}(K, N)$.

We now describe in details our results.

The Monge minimization problem

Let (X,d) be a complete and separable metric space (Polish space) and $\mathcal{P}(X)$ be the set of Borel probability measures over X. Given $\mu,\nu \in \mathcal{P}(X)$ we will study the Monge transportation problem, i.e. the minimization of the functional

$$\mathcal{I}(T) = \int_X d_L(x, T(x))\mu(dx)$$

where T varies over all Borel maps $T : X \to X$ such that $T_\sharp \mu = \nu$ and d_L is a Borel, possibly not finite, geodesic distance over X. We will present: existence results in the case of d_L non-branching, the application of our techniques to the Wiener space and then existence results also for the general case of possibly branching geodesic distance cost.

We recall briefly which are the main results concerning the existence of solutions for the Monge minimization problem, referring to the monographies [30, 31] for a deeper insight on optimal transportation. In the original formulation given by Monge the problem was settled in $\mathbb{R}^3$, with the cost given by the Euclidean norm and the measures μ,ν were supposed to be absolutely continuous and supported on two disjoint compact sets. The original problem remained unsolved for a long time. In 1978 Sudakov [28] claimed to have a solution for any distance cost function induced by a norm: an essential ingredient in the proof was that if $\mu \ll \mathcal{L}^d$ and $\mathcal{L}^d$-a.e. $\mathbb{R}^d$ can be decomposed into convex sets of dimension k, then then the conditional probabilities are absolutely continuous with respect to the $\mathcal{H}^k$ measure of the correct dimension. But it turns out that when $d > 2$, $0 < k < d-1$ the property claimed by Sudakov is not true. An example with $d = 3$, $k = 1$ can be found in [22].

The Euclidean case has been correctly solved only during the last decade. L. C. Evans and W. Gangbo in [19] solved the problem under the assumptions that $\mathrm{supp}[\mu]\cap\mathrm{supp}[\nu] = \emptyset$, $\mu,\nu \ll \mathcal{L}^d$ and their densities are Lipschitz functions with compact support. The first existence results for general absolutely continuous measures μ,ν with compact support have been independently obtained by L. Caffarelli, M. Feldman and R.J. McCann in [13] and by N. Trudinger and X.J. Wang in [29]. Afterwards M. Feldman and R.J. McCann [20] extended the results to manifolds with geodesic cost. The case of a general norm as cost function on $\mathbb{R}^d$, including also the case with non strictly convex unitary ball, has been solved first in the particular case of crystalline norm by L. Ambrosio, B. Kirchheim and A. Pratelli in [4], and then in fully generality independently by L. Caravenna in [14] and by T. Champion and L. De Pascale in [18].

The non-branching case

We will prove that given a d_L-cyclically monotone transference plan $\pi \in \Pi(\mu,\nu)$, under appropriate assumptions on the first marginal μ and on the behavior of d_L-geodesics, there exists a transport map $T : X \to X$ (recall that transport means $T_\sharp\mu = \nu$) with the same transference cost of π. Since we do not require d_L to be l.s.c., the existence of an optimal transference plan is not guaranteed and our strategy doesn't rely on a possible optimality of π. However since optimality implies d_L-cyclical monotonicity, our results include existence theorems for optimal transport maps provided the existence of an optimal transference plan. Moreover it is worth noting that due to the lack of regularity of d_N we will not use the existence of optimal potentials (ϕ,ψ).

We present our approach. The presence of 1-dimensional sets (the geodesics) along which the cost is linear is a strong degeneracy for transport problems. This degeneracy is equivalent to the following problem in $\mathbb{R}$: if μ is concentrated on $(-\infty, 0]$, and ν is concentrated on $[0,+\infty)$, then every transference plan is optimal for the 1-dimensional distance cost $|\cdot|$. In fact, every $\pi \in \Pi(\mu,\nu)$ is supported on the set $(-\infty, 0] \times [0,+\infty)$, on which $|x-y| = y-x$ and thus

$$\int |x-y|\pi(dxdy) = -\int x\mu(dx) + \int y\nu(dy).$$

Nevertheless, for this easy case an explicit map $T : \mathbb{R} \to \mathbb{R}$ can be constructed if μ is without atoms (i.e. continuous): the easiest choice is the monotone map, a minimizer of the quadratic cost $|\cdot|^2$.

The strategy suggested by the above simple case is the following:

1. reduce the problem to transportation problems along distinct geodesics;
2. show that the disintegration of the marginal μ on each geodesic is continuous;
3. find a transport map on each geodesic and piece them together.

While the last point can be seen as an application of selection principles in Polish spaces, the first two points are more subtle.

The geodesics used by a given d_L-cyclically monotone transference plan π to transport mass can be obtained from a set Γ on which π is concentrated. Under the non-branching assumption, a cyclically monotone plan π yields a natural partition R of a subset of the transport set $\mathcal{T}_e$, i.e. the set of points on the geodesics used by π: defining

- the set $\mathcal{T}$ made of inner points of geodesics,
- the set $a \cup b := \mathcal{T}_e \setminus \mathcal{T}$ of initial points a and end points b,

the non-branching assumption and the cyclical monotonicity of Γ imply that the geodesics used by π are a partition on $\mathcal{T}$. In general in a there are points from which more than geodesic starts and in b there are points in which more than one geodesic ends, hence being on a geodesic can't be an equivalence relation on the set $a \cup b$. For example one can think to the unit circle with $\mu = \delta_0$ and $\nu = \delta_\pi$.

We note here that π gives also a direction along each component of R, as the one dimensional example above shows.

Even if we have a natural partition R of $\mathcal{T}$ and $\mu(a \cup b) = 0$, the reduction to transport problems on the equivalence classes is not straightforward: a necessary and sufficient condition is that the disintegration of the measure μ is strongly consistent, that is equivalent to the existence a μ-measurable quotient map $f : \mathcal{T} \to \mathcal{T}$ of the equivalence relation R. Since this partition is closely related to the geodesics of d_L, the strong consistency will follow from topological properties of the d_L-geodesic considered as curves in (X, d): in fact we require that they are d-continuous and locally compact

Then we can write

$$m := f_\sharp \mu, \quad \mu = \int \mu_y m(dy), \quad \mu_y(f^{-1}(y)) = 1,$$

i.e. the conditional probabilities μ_y are concentrated on the counterimages $f^{-1}(y)$ (which are single geodesics). We can obtain the one dimensional problems by partitioning π w.r.t. the partition $R \times (X \times X)$,

$$\pi = \int \pi_y m(dy), \quad \nu = \int \nu_y m(dy) \quad \nu_y := (P_2)_\sharp \pi_y,$$

and considering the one dimensional problems along the geodesic $R(y)$ with marginals μ_y, ν_y and cost $|\cdot|$, the length on the geodesic.

To next step is study the continuity of the conditional probabilities μ_y and whether $\mu \llcorner \mathcal{T}_e = \mu \llcorner \mathcal{T}$ holds true. To pursue this aim we consider a natural operation on sets: the translation along geodesics. If A is a subset of $\mathcal{T}$, we denote by A_t the set translated by t in the direction determined by π. It turns out that $\mu(a \cup b) = 0$ and the continuity of μ_y both depend on how the function $t \mapsto \mu(A_t)$ behaves.

Theorem 0.1 (Lemma 1.30 and Proposition 1.31)**.** *If $\sharp\{t > 0 : \mu(A_t) > 0\}$ is uncountable for all A Borel such that $\mu(A) > 0$, then $\mu(a \cup b) = 0$ and the conditional probabilities μ_y are continuous.*

This is sufficient to solve the Monge problem, i.e. to find a transport map which has the same cost as π. A second result concerns a stronger regularity property of μ_y obtained under a slightly more restrictive assumption.

Theorem 0.2 (Theorem 1.34)**.** *Assume that $\mathcal{L}^1(\{t > 0 : \mu(A_t) > 0\}) > 0$ for all A Borel such that $\mu(A) > 0$. Then $\mu(a \cup b) = 0$ and μ_y is a.c. w.r.t. the 1-dimensional Hausdorff measure $\mathcal{H}^1_{d_L}$ induced by d_L.*

Assuming $d_L \geq d$, the assumption of the previous theorem allows to define a current in (X, d) representing the vector field corresponding to the translation $A \mapsto A_t$, and moreover to solve the equation

$$\partial U = \mu - \nu$$

is the sense of current in metric space.

The final result under the non-branching assumption is the stability of the regularity of the structures introduced so far under Measured-Gromov-Hausdorff like convergence of $(X_n, d_n, d_{L,n}, \pi_n)$. The conclusion is that a sort of uniform integrability condition on the conditional probability w.r.t. $\mathcal{H}^1_{d_{L,n}}$ passes to the limit, so that one can verify by approximation if Theorem 0.2 holds.

We also present an application of this stability result in the case $d = d_L$. Consider a reference measure $\eta \in \mathcal{P}(X)$ such that (X, d, η) is a non-branching metric measure space satisfying the $\mathsf{MCP}(K, N)$ for $K \in \mathbb{R}$ and $N \geq 1$. The stability result together with MCP condition implies that Assumption 2 holds for η w.r.t. the optimal flow induced by any d-monotone plan $\pi \in \Pi(\mu, \nu)$. Hence if $\mu \ll \eta$, the existence of a minimizer for the Monge minimization problem with marginal μ and ν follows.

Corollary 0.3. *Let K, N be real numbers with $N \geq 1$. Let (X, d, η) be a metric measure space satisfying $\mathsf{MCP}(K, N)$ and $\mu, \nu \in \mathcal{P}(X)$ with $\mu \ll \eta$. Then the Monge minimization problem between μ and ν with distance cost d admits a solution.*

The results presented are contained in a joint work with Stefano Bianchini, [9].

The Wiener Space

Let $(X, \|\cdot\|)$ be an ∞-dimensional separable Banach space, $\gamma \in \mathcal{P}(X)$ be a non degenerate Gaussian measure over X and $H(\gamma)$ be the corresponding Cameron-Martin space with Hilbertian norm $\|\cdot\|_{H(\gamma)}$. Given two probability measures $\mu, \nu \in \mathcal{P}(X)$, we will prove the existence of a solution for the following Monge minimization problem

$$\min_{T : T_\sharp \mu = \nu} \int_X \|x - T(x)\|_{H(\gamma)} \mu(dx), \tag{0.0.1}$$

provided μ and ν are both absolutely continuous w.r.t. γ.

The Wiener space $(X, \|\cdot\|, \gamma)$ fits into the framework developed for the general non-branching case, indeed: $(X, \|\cdot\|)$ is a separable Banach space, the Cameron-Martin norm $\|\cdot\|_{H(\gamma)}$ is lower semi-continuous w.r.t. $\|\cdot\|$, it is geodesic and non-branching. Hence from the theory already developed we know that we can reduce the problem to transportation problems along distinct geodesics. The main issue will be to show that the disintegration of the marginal μ on each geodesic is continuous. At that the point, as already explained before, we will have the existence of an optimal map on each geodesic. Then, gluing together all the one-dimensional optimal maps, we obtain a global optimal map.

Through a small modification of the evolution along the transport set introduced before, we prove a result similar to Theorem 0.1: if Γ is the support of a given $\|\cdot\|_{H(\gamma)}$-cyclically monotone transference plan π and $A \subset \mathcal{T}_e$, we consider the set $T_t(\Gamma \cap A \times X)$ where T_t is the map from $X \times X$ to X that associates to a couple of points its convex combination at time t.

It turns out that the fact that $\mu(a) = 0$ and the measures μ_y are continuous depends on the behavior of the function $t \mapsto \gamma(T_t(\Gamma \cap A \times X))$.

Theorem 0.4 (Proposition 2.14 and Proposition 2.15)**.** *If for every A with $\mu(A) > 0$ there exists a sequence $t_n \searrow 0$ and a positive constant C such that $\gamma(T_{t_n}(\Gamma \cap A \times X)) \geq C\mu(A)$, then $\mu(a) = 0$ and the conditional probabilities μ_y and ν_y are continuous.*

This result implies that the existence of a minimizer of the Monge problem is equivalent to the regularity properties of $t \mapsto \gamma(T_t(\Gamma \cap A \times X))$. Hence the problem is reduced to verify that the Gaussian measure γ satisfies the assumptions of Theorem 0.4.

Let $\mu = \rho_1\gamma$ and $\nu = \rho_2\gamma$ and assume that ρ_1 and ρ_2 are bounded. Then we find suitable d-dimensional measures μ_d, ν_d, absolutely continuous w.r.t. the d-dimensional Gaussian measure γ_d, converging to μ and ν respectively, such that (Theorem 2.18) γ_d verifies

$$\gamma_d(T_t(\Gamma_d \cap A \times X)) \geq C\mu_d(A)$$

where the evolution now is induced by the transport problem between μ_d and ν_d (the set Γ_d will be the graph of an optimal map between μ_d and ν_d) and the constant C does not depend on the dimension. Passing to the limit as $d \nearrow \infty$, we prove the same property for γ. Hence the existence result is proved for measures with bounded densities. To obtain the existence result in full generality we observe that the transport set $\mathcal{T}_e$ is a transport set also for transport problems between measures satisfying the uniformity condition stated above (Proposition 2.20 and Proposition 2.21).

The assumption that both μ and ν are a.c. with respect to γ is fundamental. Indeed take as example a diffuse measure μ and $\nu = \delta_x$, then the constant in the evolution estimate induced by the optimal transference plan will depend on the dimension and passing to the limit we lose all the informations on the evolution.

Theorem 0.5 (Theorem 2.22). *Let $\mu, \nu \in \mathcal{P}(X)$ with $\mu, \nu \ll \gamma$. Then there exists a solution for the Monge minimization problem*

$$\min_{T : T_\sharp\mu = \nu} \int \|x - T(x)\|_{H(\gamma)}\mu(dx).$$

Moreover we can find T invertible.

The results presented on the Monge problem in Wiener space are taken from [16].

The general case

We come back to the general analysis of the Monge minimization problem removing the non-branching assumption. To avoid mistakes, instead of d_L, we will denote the possibly branching and not finite, geodesic distance on X with d_N. We will prove that given a d_N-cyclically monotone transference plan $\pi \in \Pi(\mu, \nu)$, under appropriate assumptions on the first marginal and on the plan π, there exists an admissible map $T : X \to X$ with the same transference cost of π. As in the non-branching case, due to lack of regularity, our strategy doesn't rely on a possible optimality of π and we don't use optimal potentials (ϕ, ψ).

We will present an application of our results to optimal transportation in $\mathbb{R}^d$ around a convex smooth obstacle (the obstacle problem).

The steps to solve the Monge problem with branching distance cost are the following:

1. reduce the problem, via Disintegration Theorem, to transportation problems in sets where, under a regularity assumption on the first marginal and on π, we know how to produce an optimal map;

2. show that the disintegration of the first marginal μ on each of this sets verifies this regularity assumption;

3. find a transport map on each of these sets and piece them together.

In the easier case of d_N non-branching, given a d_N-cyclically monotone transference plan it is always possible to reduce the problem on single geodesics. The reduced problem becomes essentially one dimensional and there the precise regularity assumption is that the first marginal has no atoms (is continuous). Now this reduction can't be done anymore and there is not another reference set where the existence of Monge minimizer is known.

The reduction set will be a concatenation of more geodesics and to produce an optimal map we will need a regularity assumption also on the shape of this set.

As in the non-branching case, starting from from a d_N-cyclical monotone set Γ on which π is concentrated one can construct the set of transport rays R, the transport set $\mathcal{T}_e$, i.e. the set of geodesics used by π, and from them construct

- the set $\mathcal{T}$ made of inner points of geodesics,
- the set $a \cup b := \mathcal{T}_e \setminus \mathcal{T}$ of initial points a and end points b.

Since branching of geodesics is admitted, R is not a partition on $\mathcal{T}$. To obtain an equivalence relation we have to consider the set H of chain of transport rays: it is the set of couples (x, y) such that we can go from x to y with a finite number of transport rays such that their common points are not final or initial points. Hence H will provide the partition of the transport set $\mathcal{T}$ and each equivalence class, $H(y)$ for y in the quotient space, will be a reduction set.

As already explained, to perform a real reduction to transport problems on the equivalence classes we also need the disintegration of μ w.r.t. the partition H to be strongly consistent. This is equivalent to the fact that there exists a μ-measurable quotient map $f : \mathcal{T} \to \mathcal{T}$ of the equivalence relation induced by the partition.
Since this partition is closely related to the geodesics of d_N, the strong consistency will follow from a topological property of the geodesic as set in (X, d) and from a metric property of d_N as a function:

(1.a) each chain of transport rays $H(y)$ restricted to a d_N closed ball is d-closed;

(1.b) $d_N(x, \cdot)$ restricted to $H(x)$ is bounded on d-bounded sets.

Observe that these conditions on H and d_N are the direct generalization of the ones on geodesics considered in the non-branching case (continuity and local compactness) and they depend on the particular choice of the transference plan. This assumptions permit to disintegrate μ restricted to $\mathcal{T}$. Hence one can write

$$\mu \llcorner_{\mathcal{T}} = \int \mu_y m(dy), \quad m := f_\sharp \mu, \quad \mu_y(f^{-1}(y)) = 1,$$

i.e. the conditional probabilities μ_y are concentrated on the counterimages $f^{-1}(y)$ (which is an equivalence class). The reduced problems are obtained by disintegrating π w.r.t. the partition $H \times (X \times X)$,

$$\pi \llcorner_{\mathcal{T} \times \mathcal{T}} = \int \pi_y m(dy), \quad \nu = \int \nu_y m(dy) \quad \nu_y := (P_2)_\sharp \pi_y,$$

and considering the problems on the sets $H(y)$ with marginals μ_y, ν_y and cost d_N.

To next step is study the continuity of the conditional probabilities μ_y and whether $\mu \llcorner_{\mathcal{T}_e} = \mu \llcorner_{\mathcal{T}}$ holds true. To pursue this aim we consider a natural operation on sets: the translation along geodesics. If A is a subset of $\mathcal{T}$, we denote by A_t the set translated by t in the direction determined by π. A rigorous definition of the translation of sets along geodesic will be given during the chapter. It turns out that $\mu(a \cup b) = 0$ and the continuity of μ_y both depend again on how the function $t \mapsto \mu(A_t)$ behaves. Indeed assuming that:

(2) for all A Borel there exists a sequence $\{t_n\} \subset \mathbb{R}$ and $C > 0$ such that $\mu(A_{t_n}) \geq C\mu(A)$ as $t_n \to 0$,

we have the following.

Theorem 0.6 (Proposition 3.16 and Proposition 3.17)**.** *If Assumption (2) holds, then $\mu(a \cup b) = 0$ and the conditional probabilities μ_y are continuous.*

At this level of generality we don't know how to obtain a d_N-monotone admissible map for the restricted problem even if the marginals μ_y satisfy some regularity assumptions. Therefore we need to assume that $H(y)$ has a particular structure:

(3) for m-a.e. y, the chain of transport rays $H(y)$ is contained, up to set of μ_y-measure zero, in an uncountable "increasing" family of measurable sets.

A rigorous formulation of Assumption (3) and of "increasing" will be given later on. If $H(y)$ satisfies Assumption (3), then we can perform a disintegration of μ_y with respect to the partition induced by the uncountable "increasing" family of sets. Then if the quotient measure and the marginal measures of μ_y are continuous, we prove the existence of an optimal map between μ_y and ν_y.

Theorem 0.7 (Proposition 3.19 and Theorem 3.18). *Let $\pi \in \Pi(\mu, \nu)$ be a d_N-monotone plan concentrated on a set Γ. Assume that Assumptions (1.a), (1.b), (2), (3) holds and that the quotient measure and the marginal measures of μ_y are continuous for m-a.e. y. Then there exists an admissible map with the same transference cost of π.*

It follows immediately that if we also assume that π is optimal in the hypothesis of Theorem 0.7, then the Monge minimization problem admits a solution.

Before presenting an application of Theorem 0.7, we prefer to do a brief summary of the theoretical results obtained on this general case. Let $\pi \in \Pi(\mu, \nu)$ be a d_N-cyclically monotone transference plan concentrated on a set Γ. We consider the corresponding family of chain of transport rays and, if assumptions (1.a) and (1.b) are satisfied, we can perform, neglecting the set of initial points, a disintegration of μ, ν and π with respect to the partition induced by the chain of transport rays. Then if assumption (2) is satisfied it follows that the set of initial points is μ-negligible and the conditional probabilities μ_y are continuous. Since the geometry of $H(y)$ can be wild, we need another assumption to build a d_N-monotone transference map between μ_y and ν_y. If $H(y)$ satisfies assumption (3) we can perform another disintegration and, under additional regularity of the conditional probabilities of μ_y and of the quotient measure of μ_y, we prove the existence of a d_N-monotone transference map between μ_y and ν_y. Applying the same reasoning for m-a.e. y we prove the existence of a transport map T between μ and ν that has the same transference cost of the given d_N-cyclically monotone plan π.

We now pass to an application of Theorem 0.7. Consider a hyper-surface $M \subset \mathbb{R}^d$ that is the boundary of a convex and compact set C. Let X be the closure, in the euclidean topology, of $\mathbb{R}^d \setminus C$ and take as cost function d_M: the minimum of the euclidean length among all Lipschitz curves in X. Hence C will play the role of a convex smooth obstacle. We will study the Monge minimization problem in X with d_M as transport cost and we will prove that if μ is absolutely continuous w.r.t. $\mathcal{L}^d$, the Monge minimization problem admits a solution.

It is worth noting that the hypothesis of Theorem 0.7, namely Assumptions (1.a), (1.b), (2), (3), are all about the behavior of a given d_N-cyclically monotone transference plan. For the obstacle problem we will prove that any d_M-cyclically monotone transference plan satisfies this hypothesis. Therefore this concrete example provide also a confirmation of the validity of the proposed strategy and, in particular, of the non artificiality of the proposed assumptions.

The results of this section are taken from [15].

A partial globalization result for curved metric measure spaces

Analysis on singular spaces is one big challenge in mathematics. An important class of singular spaces is the class of metric measure spaces with generalized lower bounds on the Ricci curvature formulated in terms of optimal transportation. The condition of lower bounds on Ricci curvature for singular spaces and the corresponding class of spaces have been introduced by Sturm in [26, 27] and independently by Lott and Villani in [23].

This condition called *curvature-dimension condition* $\mathsf{CD}(K, N)$ depends on two parameters K and N, playing the role of a curvature and dimension bound, respectively. We recall two important properties of the condition $\mathsf{CD}(K, N)$:

- the curvature-dimension condition is stable under convergence of metric measure spaces with respect to the L^2-transportation distance $\mathbb{D}$ introduced in [26];
- a complete Riemannian manifold satisfies $\mathsf{CD}(K, N)$ if and only if its Ricci curvature is bounded from below by K and its dimension from above by N.

Moreover a broad variety of geometric and functional analytic properties can be deduced from the curvature-dimension condition $\mathsf{CD}(K, N)$: Brunn-Mikowski inequality, Bishop-Gromov volume comparison theorem, Bonnet-Myers theorem, the doubling property and local Poincaré inequalities on balls. All this listed results are quantitative results (volume of intermediate points, volume growth, upper bound on the diameter and so on) depending on K, N.

There is a weak variant of $\mathsf{CD}(K,N)$, the *measure-contraction property* $\mathsf{MCP}(K,N)$, introduced in [24] and [27]. In the setting of non-branching metric measure spaces it is proven that condition $\mathsf{CD}(K,N)$ implies $\mathsf{MCP}(K,N)$. Roughly spoken, $\mathsf{CD}(K,N)$ is a condition on the optimal transport between any pair of absolutely continuous probability measure on M, whereas $\mathsf{MCP}(K,N)$ is a condition on the optimal transport between Dirac masses and the uniform distribution m on M. Nevertheless a great part of the geometric and functional analytic properties verified by spaces satisfying the condition $\mathsf{CD}(K,N)$ are also verified by spaces satisfying the $\mathsf{MCP}(K,N)$:

- generalized Bishop-Gromov volume growth inequality;
- doubling property;
- a bound on the Hausdorff dimension;
- generalized Bonnet-Myers theorem,

and many other. Again this results are in a quantitative form depending on K, N. For a complete list of analytic consequences of the measure contraction property see [27].

Among the relevant questions on $\mathsf{CD}(K,N)$ that are still open, we are interested in studying the following one: can we say that a metric measure space (M,d,m) satisfies $\mathsf{CD}(K,N)$ provided $\mathsf{CD}(K,N)$ holds true locally on a family of sets M_i covering M?
In other words it is still not known whether $\mathsf{CD}(K,N)$ verifies the globalization property (or the local-to-global property).

A partial answer to this problem is contained in the work by Bacher and Sturm [7]: they proved that if a metric measure space (M,d,m) verifies the local curvature-dimension condition $\mathsf{CD}_{loc}(K,N)$ then it verifies the global reduced curvature-dimension condition $\mathsf{CD}^*(K,N)$. The latter is strictly weaker than $\mathsf{CD}(K,N)$ and a converse implication can be obtained only changing the value of the lower bound on the curvature: condition $\mathsf{CD}^*(K,N)$ implies $\mathsf{CD}(K^*,N)$ where $K^* = K(N-1)/N$. Therefore $\mathsf{CD}^*(K,N)$ gives worse geometric and analytic information than $\mathsf{CD}(K,N)$.

Our contribution towards a complete answer to the previous question is the following: we prove that if (M,d,m) is a non-branching metric measure space that verifies $\mathsf{CD}_{loc}(K,N)$ then (M,d,m) verifies $\mathsf{MCP}(K,N)$.

Hence our result implies that from the local condition $\mathsf{CD}_{loc}(K,N)$ one can obtain all the global geometric and functional analytic consequences implied by $\mathsf{MCP}(K,N)$ and therefore the geometric and functional analytic consequences are obtained in the sharp quantitative version.

We now present our approach to the problem.
As already pointed out, the curvature-dimension condition $\mathsf{CD}(K,N)$ prescribes how the volume of a given set is affected by curvature when it is moved via optimal transportation. Condition $\mathsf{CD}(K,N)$ impose that the distortion is ruled by the coefficient $\tau_{K,N}^{(t)}(\theta)$ depending on the curvature K, on the dimension N, on the time of the evolution t and on the point θ.

The main feature of the coefficient $\tau_{K,N}^{(t)}(\theta)$ is that it is obtained mixing two different information on how the volume should evolve: an $(N-1)$-dimensional distortion depending on the curvature K by and a one dimensional evolution that doesn't feel the curvature. To be more precise

$$\tau_{K,N}^{(t)}(\theta) = t^{1/N}\sigma_{K,N-1}^{(t)}(\theta)^{(N-1)/N},$$

where $\sigma_{K,N-1}^{(t)}(\theta)^{(N-1)/N}$ contains the information on the $(N-1)$-dimensional volume distortion and the evolution in the remaining direction is ruled just by $t^{1/N}$. This is a clear similarity with the Riemannian case.

Our aim is, starting from $\mathsf{CD}_{loc}(K,N)$, to isolate a local $(N-1)$-dimensional condition ruled by the coefficient $\sigma_{K,N-1}^{(t)}(\theta)$ and then, using the easier structure of $\sigma_{K,N-1}^{(t)}(\theta)$, obtain a global $(N-1)$-dimensional condition with coefficient $\sigma_{K,N-1}^{(t)}(\theta)$. At that point, using Hölder inequality and the fact that the missing direction is not affected by curvature, it is not difficult to pass from the $(N-1)$-dimensional version to the full-dimensional version with coefficient $\tau_{K,N}^{(t)}(\theta)$.

However to detect a local $(N-1)$-dimensional condition it is necessary to decompose the whole evolution into a family of $(N-1)$-dimensional evolutions. Considering the optimal transport between a Dirac mass in x_0 and the uniform distribution m, the family of spheres around x_0 immediately provides the correct $(N-1)$-dimensional evolutions. This motivates why we obtain $\mathsf{MCP}(K,N)$ and not $\mathsf{CD}(K,N)$.

We state the main result of this part of the thesis.

Theorem 0.8 (Theorem 4.16)**.** *Let (M,d,m) be a non-branching metric measure space. Assume that (M,d,m) satisfies $\mathsf{CD}_{loc}(K,N)$. Then (M,d,m) satisfies $\mathsf{MCP}(K,N)$.*

We tried to make notation as much unified as possible. Nevertheless, the main specific notations will be introduced chapter by chapter.

Chapter 1

The Monge problem for non-branching geodesic distance cost

This chapter concerns the Monge transportation problem in geodesic spaces, i.e. metric spaces with a geodesic structure. Given two Borel probability measure $\mu, \nu \in \mathcal{P}(X)$, where (X, d) is a Polish space, we study the minimization of the functional

$$\mathcal{I}(T) = \int d_L(x, T(x))\mu(dy)$$

where T varies over all Borel maps $T : X \to X$ such that $T_\sharp \mu = \nu$ and d_L is a Borel distance that makes (X, d_L) a non branching geodesic space.

Chapter 1 is organized as follows.

In Section 1.1 we define the geodesic structure (X, d, d_L) which is studied in this chapter. Section 1.2 shows how using only the d_L-cyclical monotonicity of a set Γ we can obtain a partial order relation $G \subset X \times X$ as follows (Lemma 1.10 and Proposition 1.14): xGy iff there exists $(w, z) \in \Gamma$ and a geodesic $\gamma : [0, 1] \to X$, with $\gamma(0) = w$, $\gamma(1) = z$, such that x, y belongs to γ and $\gamma^{-1}(x) \leq \gamma^{-1}(y)$. This set G is analytic, and allows to define

- the transport ray set R (1.2.4),
- the transport sets $\mathcal{T}_e$, $\mathcal{T}$ (with and without and points) (1.2.5),
- the set of initial points a and final points b (1.2.8).

Moreover we show that $R \llcorner_{\mathcal{T} \times \mathcal{T}}$ is an equivalence relation (Proposition 1.14), we can assume that the set of final points b can be taken μ-negligible (Lemma 1.18), and in two final remarks we study what happens in the case more regularity on the cost d_L is assumed, Remark 1.19 and Remark 1.20.
Notice that in the case $d = d_L$ the existence of a Lipschitz potential φ, one can take

$$\Gamma = G = \Big\{(x, y) : \varphi(x) - \varphi(y) = d(x, y)\Big\}.$$

Thus the main result of this section is that these sets can be defined even if the potential does not exist.

Section 1.3 proves that the continuity and local compactness of geodesics imply that the disintegration induced by R on $\mathcal{T}$ is strongly consistent (Proposition 1.24): as Example 1 shows, the strong consistency of the disintegration is a non trivial property of the metric spaces we are considering.
Using this fact, we can define an order preserving map g which maps our transport problem into a transport problem on $\mathcal{S} \times \mathbb{R}$, where $\mathcal{S}$ is a cross section of R (Proposition 1.26). Finally we show that

This chapter is based on the joint work with Stefano Bianchini [9].

under this assumption there exists a transference plan with the same cost of π which leaves the common mass $\mu \wedge \nu$ at the same place (note that in general this operation lowers the transference cost).

In Section 1.4 we prove Theorem 0.1 and Theorem 0.2. We first introduce the operation $A \mapsto A_t$, the translation along geodesics (1.4.1), and show that $t \mapsto \mu(A_t)$ is a Souslin function if A is analytic (Lemma 1.29).
Next we show that under the assumption

$$\mu(A) > 0 \quad \Longrightarrow \quad \sharp\big\{t > 0 : \mu(A_t) > 0\big\} > \aleph_0$$

the set of initial points a is μ-negligible (Lemma 1.30) and the conditional probabilities μ_y are continuous. Finally, we show that under the stronger assumption

$$\mu(A) > 0 \quad \Longrightarrow \quad \int_{\mathbb{R}^+} \mu(A_t) dt > 0, \tag{1.0.1}$$

the conditional probabilities μ_y are a.c. w.r.t. $\mathcal{H}^1_{d_L}$ (Theorem 1.34). A final result shows that actually Condition (1.0.1) yields that $t \mapsto \mu(A_t)$ has more regularity than just integrability (Proposition 1.35) it is in fact continuous

After the above results, the solution of the Monge problem is routine, and it is done in Theorem 1.37 of Section 1.5.

Under Condition 1.0.1 and $d \leq d_L$, in Section 1.6 we give a dynamic interpretation to the transport along geodesics. In Definition 1.39 we define the current $\dot{g}$ in (X, d), which represents the flow induced by the transference plan π. Not much can be said of this flow, unless some regularity assumptions are considered. These assumptions are the natural extensions of properties of transportation problems in finite dimensional spaces.
If there exists a background measure η whose disintegration along geodesics satisfies

$$\eta = \int q_y \mathcal{H}^1_{d_L} m(dy), \quad q_y \in \mathrm{BV}, \int \mathrm{Tot.Var.}(q_y) m(dy) < +\infty,$$

then $\dot{g}$ is a normal current, i.e. its boundary is a bounded measure on X (Lemma 1.40).
We can also consider the problem $\partial U = \mu - \nu$ in the sense of currents: Proposition 1.42 gives a solution, and in the case $q_y(t) > 0$ for $\mathcal{H}^1_{d_L}$-a.e. t we can write represent $U = \rho \dot{g}$, i.e. the flow $\dot{g}$ multiplied by a scalar density ρ (Corollary 1.44).

In Section 1.7 we address the stability of the assumptions under Measure-Gromov-Hausdorff-like convergence of structures $(X_n, d_n, d_{L,n}, \pi_n)$. Under a uniform integrability condition of $\mu_{y,n}$ w.r.t. $\mathcal{H}^1_{d_{L,n}}$ and a uniform bound on the π_n transportation cost (Assumption 4 of Section 1.7.2), we show that the marginal μ can be represented as the image of a measure $rm \otimes \mathcal{L}^1$ by a Borel function $h : \mathcal{T} \times \mathbb{R} \to \mathcal{T}_e$, with $r \in L^1(m \otimes \mathcal{L}^1)$ (Proposition 1.57). The key feature of h is that $t \mapsto h(y, t)$ is a geodesic of $\mathcal{T}$ for m-a.e. $y \in \mathcal{T}$.
Thus while $h(0, \mathcal{T})$ is not a cross section for R (in that case we would have finished the proof), in Proposition 1.47 we show which conditions on h imply that μ can be disintegrated with a.c. conditional probabilities, and we verify that this is our case in Theorem 1.58.
In two remarks we suggest how to pass also uniform estimates on the disintegration on $(X_n, d_n, d_{L,n})$ to the transference problem in (X, d, d_L) (Remark 1.48 and Remark 1.59).

In Section 1.8 we consider an application of the results obtained in the previous sections. We assume $d = d_L$ and the existence of background probability measure η such that (X, d, η) satisfies $MCP(K, N)$ (Definition 1.60). In this framework we prove that for any d-cyclically monotone transference plan π, η admits a disintegration along the geodesics used by π with marginal probabilities absolutely continuous w.r.t. $\mathcal{H}^1$ (Theorem 1.64). This implies directly (Corollary 1.65) that if $\mu \ll \eta$ the Monge minimization problem with marginals μ and ν admits a solution. The final result of the section (Lemma 1.66) shows that we can solve the dynamical problem $\partial U = \mu - \nu$ with $U = \rho \dot{g}$, and if the support of μ and ν are disjoint U is a normal current.

The last section contains two important examples. In Example 1 we show that if the geodesics are not locally compact, then in general the disintegration along transport rays is not strongly supported. In Example 2 we show that under our assumptions the c-monotonicity is not sufficient for optimality.

1.1 Metric setting

In this section we recall some general facts about geodesic spaces and we refer to [12].

Definition 1.1. A *length structure* on a topological space X is a class $\mathtt{A}$ of admissible paths, which is a subset of all continuous paths in X, together with a map $L : \mathtt{A} \to [0, +\infty]$: the map L is called *length of path*. The class $\mathtt{A}$ satisfies the following assumptions:

closure under restrictions if $\gamma : [a, b] \to X$ is admissible and $a \leq c \leq d \leq b$, then $\gamma\llcorner_{[c,d]}$ is also admissible.

closure under concatenations of paths if $\gamma : [a, b] \to X$ is such that its restrictions γ_1, γ_2 to $[a, c]$ and $[c, b]$ are both admissible, then so is γ.

closure under admissible reparametrizations for an admissible path $\gamma : [a, b] \to X$ and a for $\varphi : [c, d] \to [a, b]$, $\varphi \in B$, with B class of admissible homeomorphisms that includes the linear one, the composition $\gamma(\varphi(t))$ is also admissible.

The map L satisfies the following properties:

additivity $L(\gamma\llcorner_{[a,b]}) = L(\gamma\llcorner_{[a,c]}) + L(\gamma\llcorner_{[c,b]})$ for any $c \in [a, b]$.

continuity $L(\gamma\llcorner_{[a,t]})$ is a continuous function of t.

invariance The length is invariant under admissible reparametrizations.

topology Length structure agrees with the topology of X in the following sense: for a neighborhood U_x of a point $x \in X$, the length of paths connecting x with points of the complement of U_x is separated from zero:

$$\inf \big\{ L(\gamma) : \gamma(a) = x, \gamma(b) \in X \setminus U_x \big\} > 0.$$

Given a length structure, we can define a distance

$$d_L(x, y) = \inf \Big\{ L(\gamma) : \gamma : [a, b] \to X, \gamma \in \mathtt{A}, \gamma(a) = x, \gamma(b) = y \Big\},$$

that makes (X, d_L) a metric space (allowing d_L to be $+\infty$). The metric d_L is called *intrinsic*.

It follows from Proposition 2.5.9 of [12] that every admissible curve of finite length admits a constant speed parametrization, i.e. γ defined on $[0, 1]$ and $L(\gamma\llcorner[t, t']) = v(t' - t)$, with v velocity.

Definition 1.2. A length structure is said to be *complete* if for every two points x, y there exists an admissible path joining them whose length $L(\gamma)$ is equal to $d_L(x, y)$.

In other words, a length structure is complete if there exists a shortest path between two points.

Intrinsic metrics associated with complete length structure are said to be *strictly intrinsic*. The metric space (X, d_L) with d_L strictly intrinsic is called a *geodesic space*. A curve whose length equals the distance between its end points is called *geodesic*.

Definition 1.3. Let (X, d_L) be a metric space. The distance d_L is said to be *strictly convex* if, for all $r \geq 0$, $d_L(x, y) = r/2$ implies that

$$\{z : d_L(x, z) = r\} \cap \{z : d_L(y, z) = r/2\}$$

is a singleton.

The definition can be restated in geodesics spaces as: geodesics cannot bifurcate in the interior, i.e. *the geodesic space* (X, d_L) *is not branching.* An equivalent requirement is that if $\gamma_1 \neq \gamma_2$ and $\gamma_1(0) = \gamma_2(0)$, $\gamma_1(1) = \gamma_2(1)$, then $\gamma_1((0, 1)) \cap \gamma_2((0, 1)) = \emptyset$ and such geodesics do not admit a geodesic extension i.e. they are not a part of a longer geodesic.

From now on we assume the following:

1. (X, d) Polish space;
2. $d_L : X \times X \to [0, +\infty]$ Borel distance;
3. (X, d_L) is a non-branching geodesic space;
4. geodesics are continuous w.r.t. d;
5. geodesics are locally compact in (X, d): if γ is a geodesic for (X, d_L), then for each $x \in \gamma$ there exists r such that $\gamma^{-1}(\bar{B}_r(x))$ is compact in $\mathbb{R}$.

Since we have two metric structures on X, we denote the quantities relating to d_L with the subscript L: for example

$$B_r(x) = \big\{y : d(x, y) < r\big\}, \quad B_{r,L}(x) = \big\{y : d_L(x, y) < r\big\}.$$

In particular we will use the notation

$$D_L(x) = \big\{y : d_L(x, y) < +\infty\big\},$$

$(\mathcal{K}, d_H)$ for the compact sets of (X, d) with the Hausdorff distance d_H and $(\mathcal{K}_L, d_{H,L})$ for the compact sets of (X, d_L) with the Hausdorff distance $d_{H,L}$. We recall that $(\mathcal{K}, d_H)$ is Polish.

We write

$$\gamma_{[x,y]} := \Big\{\gamma \in \mathrm{Lip}_{d_L}([0, 1]; X) : \gamma(0) = x, \gamma(1) = y, L(\gamma) = d_L(x, y)\Big\}. \tag{1.1.1}$$

With a slight abuse of notation, we will write

$$\gamma_{(x,y)} = \bigcup_{\gamma \in \gamma_{[x,y]}} \gamma((0, 1)), \quad \gamma_{[x,y]} = \bigcup_{\gamma \in \gamma_{[x,y]}} \gamma([0, 1]). \tag{1.1.2}$$

We will also use the following definition.

Definition 1.4. We say that $A \subset X$ is *geodesically convex* if for all $x, y \in A$ the minimizing geodesic $\gamma_{[x,y]}$ between x and y is contained in A:

$$\Big\{\gamma((0, 1)) : \gamma(0) = x, \gamma(1) = y, L(\gamma) = d(x, y), x, y \in A\Big\} \subset A.$$

Lemma 1.5. *If A is analytic in (X, d), then $\{x : d_L(A, x) < \epsilon\}$ is analytic for all $\epsilon > 0$.*

Proof. Observe that

$$\big\{x : d_L(A, x) < \epsilon\big\} = P_1\Big(X \times A \cap \big\{(x, y) : d_L(x, y) < \epsilon\big\}\Big),$$

so that the conclusion follows from the invariance of the class Σ^1_1 w.r.t. projections. $\square$

In particular, $\overline{A}^{d_L}$, the closure of A w.r.t. d_L, is analytic if A is analytic.

Remark 1.6. During the chapter, whenever more regularity is required, we will assume also the following hypothesis:

(2') $d_L : X \times X \to [0, +\infty]$ l.s.c. distance,

(4') $d_L(x, y) \geq d(x, y)$,

(5') $\cup_{x \in K_1, y \in K_2} \gamma_{[x,y]}$ is d-compact if K_1, K_2 are d-compact, $d_{L \llcorner K_1 \times K_2}$ uniformly bounded.

A simple computation shows that $d_L(x, y) \geq d(x, y)$ implies the following

1. d_L-compact sets are d-compact;

2. d-Lipschitz functions are d_L-Lipschitz with the same constant.

An application of Theorem 5.5, in the setting of Remark 1.6, gives a Borel function which selects a single geodesic $\gamma \in \gamma_{[x,y]}$ for any couple (x, y).

Lemma 1.7. *Assume that d_L is l.s.c.. Then there exists a Borel function $\Upsilon : X \times X \to \mathrm{Lip}_d([0,1], X)$ such that up to reparametrization $\Upsilon(x,y) \in \gamma_{[x,y]}$.*

Proof. Let

$$\begin{array}{rcl} F : X \times X & \to & \mathrm{Lip}_d([0,1], X) \\ (x,y) & \mapsto & \gamma_{[x,y]} \end{array}$$

with $\mathrm{Lip}_d([0,1], X)$ endowed with the uniform topology and $\gamma_{[x,y]}$ defined in (1.1.1).

The result follows by Theorem 5.5 observing that graph(F) is the set

$$\Big\{(x,y,\gamma) \in X \times X \times \mathrm{Lip}_d([0,1], X), L(\gamma) = d_L(x,y)\Big\}.$$

which is Borel by the l.s.c. of the map $\gamma \mapsto L(\gamma)$, and this is implied by the l.s.c. of d_L. $\square$

1.2 Optimal transportation in geodesic spaces

Let $\mu, \nu \in \mathcal{P}(X)$ and consider the transportation problem with cost $c(x,y) = d_L(x,y)$, and let $\pi \in \Pi(\mu,\nu)$ be a d_L-cyclically monotone transference plan with finite cost. By inner regularity, we can assume that the optimal transference plan is concentrated on a σ-compact d_L-cyclically monotone set $\Gamma \subset \{d_L(x,y) < +\infty\}$. By Lusin Theorem, we can require also that $d_{L \llcorner \Gamma}$ is σ-continuous:

$$\Gamma = \cup_n \Gamma_n, \ \Gamma_n \subset \Gamma_{n+1} \text{ compact}, \quad d_{L \llcorner \Gamma_n} \text{ continuous.} \tag{1.2.1}$$

In this section, using only the d_L-cyclical monotonicity of Γ, we obtain a partial order relation $G \subset X \times X$. The set G is analytic, and allows to define the transport ray set R, the transport sets $\mathcal{T}_e$, $\mathcal{T}$, and the set of initial points a and final points b. Moreover we show that $R_{\llcorner \mathcal{T} \times \mathcal{T}}$ is an equivalence relation and that we can assume the set of final points b to be μ-negligible.

Consider the set

$$\begin{aligned} \Gamma' := \bigg\{(x,y) : \exists I \in \mathbb{N}_0, (w_i, z_i) \in \Gamma \text{ for } i = 0, \dots, I, \ z_I = y \\ w_{I+1} = w_0 = x, \ \sum_{i=0}^{I} d_L(w_{i+1}, z_i) - d_L(w_i, z_i) = 0\bigg\}. \end{aligned} \tag{1.2.2}$$

In other words, we concatenate points $(x,z), (w,y) \in \Gamma$ if they are initial and final point of a cycle with total cost 0.

Lemma 1.8. *The following holds:*

1. *$\Gamma \subset \Gamma' \subset \{d_L(x,y) < +\infty\}$;*

2. *if Γ is analytic, so is Γ';*

3. *if Γ is d_L-cyclically monotone, so is Γ'.*

Proof. For the first point, set $I = 0$ and $(w_{n,0}, z_{n,0}) = (x,y)$ for the first inclusion. If $d_L(x,y) = +\infty$, then $(x,y) \notin \Gamma$ and all finite set of points in Γ are bounded.

For the second point, observe that

$$
\begin{aligned}
\Gamma' &= \bigcup_{I \in \mathbb{N}_0} P_{12}(A_I) \\
&= \bigcup_{I \in \mathbb{N}_0} P_{12}\bigg(\prod_{i=0}^{I} \Gamma \cap \bigg\{ \prod_{i=1}^{I} (w_i, z_i) : \sum_{i=0}^{I} d_L(w_{i+1}, z_i) - d_L(w_i, z_i) = 0, w_{I+1} = w_0 \bigg\} \bigg).
\end{aligned}
$$

For each $I \in \mathbb{N}_0$, since d_L is Borel, it follows that

$$
\bigg\{ \prod_{i=1}^{I} (w_i, z_i) : \sum_{i=0}^{I} d_L(w_{i+1}, z_i) - d_L(w_i, z_i) = 0, w_{I+1} = w_0 \bigg\}
$$

is Borel in $\prod_{i=0}^{I}(X \times X)$, so that for Γ analytic each set $A_{n,I}$ is analytic. Hence $P_{12}(A_I)$ is analytic, and since the class Σ^1_1 is closed under countable unions and intersections it follows that Γ' is analytic.

For the third point, observe that for all $(x_j, y_j) \in \Gamma'$, $j = 0, \ldots, J$, there are $(w_{j,i}, z_{j,i}) \in \Gamma$, $i = 0, \ldots, I_j$, such that

$$
d_L(x_j, y_j) + \sum_{i=0}^{I_j - 1} d_L(w_{j,i+1}, z_{j,i}) - \sum_{i=0}^{I_j} d_L(w_{j,i}, z_{j,i}) = 0.
$$

Hence we can write for $x_{J+1} = x_0$, $w_{j,I_j+1} = w_{j+1,0}$, $w_{J+1,0} = w_{0,0}$

$$
\sum_{j=0}^{J} d_L(x_{j+1}, y_j) - d_L(x_j, y_j) = \sum_{j=0}^{J} \sum_{i=0}^{I_j} d_L(w_{j,i+1}, z_{j,i}) - d_L(w_{j,i}, z_{j,i}) \geq 0,
$$

using the d_L-cyclical monotonicity of Γ. $\square$

Definition 1.9 (Transport rays)**.** Define the *set of oriented transport rays*

$$
G := \Big\{ (x, y) : \exists (w, z) \in \Gamma', d_L(w, x) + d_L(x, y) + d_L(y, z) = d_L(w, z) \Big\}. \tag{1.2.3}
$$

For $x \in X$, the *outgoing transport rays from x* is the set $G(x)$ and the *incoming transport rays in x* is the set $G^{-1}(x)$. Define the *set of transport rays* as the set

$$
R := G \cup G^{-1}. \tag{1.2.4}
$$

Lemma 1.10. *The following holds:*

1. *G is d_L-cyclically monotone;*
2. *$\Gamma' \subset G \subset \{d_L(x, y) < +\infty\}$;*
3. *the sets G, $R := G \cup G^{-1}$ are analytic.*

Proof. The second point follows by the definition: if $(x, y) \in \Gamma'$, just take $(w, z) = (x, y)$ in the r.h.s. of (1.2.3).

The third point is consequence of the fact that

$$
G = P_{34}\Big(\big(\Gamma' \times X \times X\big) \cap \Big\{ (w, z, x, y) : d_L(w, x) + d_L(x, y) + d_L(y, z) = d_L(w, z) \Big\} \Big),
$$

and the result follows from the properties of analytic sets.

The first point follows from the following observation: if $(x_i, y_i) \in \gamma_{[w_i,z_i]}$, then from triangle inequality

$$\begin{aligned} d_L(x_{i+1}, y_i) - d_L(x_i, y_i) + d_L(x_i, y_{i-1}) &\geq d_L(x_{i+1}, z_i) - d_L(z_i, y_i) - d_L(x_i, y_i) + d_L(x_i, y_{i-1}) \\ &= d_L(x_{i+1}, z_i) - d_L(x_i, z_i) + d_L(x_i, y_{i-1}) \\ &\geq d_L(x_{i+1}, z_i) - d_L(x_i, z_i) + d_L(w_i, y_{i-1}) - d_L(w_i, x_i) \\ &= d_L(x_{i+1}, z_i) - d_L(w_i, z_i) + d_L(w_i, y_{i-1}). \end{aligned}$$

Repeating the above inequality finitely many times one obtain

$$\sum_i d_L(x_{i+1}, y_i) - d_L(x_i, y_i) \geq \sum_i d_L(w_{i+1}, z_i) - d_L(w_i, z_i) \geq 0.$$

Hence the set G is d_L-cyclically monotone. □

Definition 1.11. Define the *transport sets*

$$\mathcal{T} := P_1\big(\mathrm{graph}(G^{-1}) \setminus \{x = y\}\big) \cap P_1\big(\mathrm{graph}(G) \setminus \{x = y\}\big), \tag{1.2.5a}$$
$$\mathcal{T}_e := P_1\big(\mathrm{graph}(G^{-1}) \setminus \{x = y\}\big) \cup P_1\big(\mathrm{graph}(G) \setminus \{x = y\}\big). \tag{1.2.5b}$$

From the definition of G it is fairly easy to prove that $\mathcal{T}$, $\mathcal{T}_e$ are analytic sets. The subscript e refers to the endpoints of the geodesics: clearly we have

$$\mathcal{T}_e = P_1(R \setminus \{x = y\}). \tag{1.2.6}$$

The following lemma shows that we have only to study the Monge problem in $\mathcal{T}_e$.

Lemma 1.12. *It holds $\pi(\mathcal{T}_e \times \mathcal{T}_e \cup \{x = y\}) = 1$.*

Proof. If $x \in P_1(\Gamma \setminus \{x = y\})$, then $x \in G^{-1}(y) \setminus \{y\}$ for some $y \in X$. Similarly, $y \in P_2(\Gamma \setminus \{x = y\})$ implies that $y \in G(x) \setminus \{x\}$ for some $x \in X$. Hence $\Gamma \setminus \mathcal{T}_e \times \mathcal{T}_e \subset \{x = y\}$. □

As a consequence, $\mu(\mathcal{T}_e) = \nu(\mathcal{T}_e)$ and any maps T such that for $\nu \llcorner_{\mathcal{T}_e} = T_\sharp \mu \llcorner_{\mathcal{T}_e}$ can be extended to a map T' such that $\nu = T_\sharp \mu$ with the same cost by setting

$$T'(x) = \begin{cases} T(x) & x \in \mathcal{T}_e \\ x & x \notin \mathcal{T}_e \end{cases} \tag{1.2.7}$$

We now use the non branching assumption.

Lemma 1.13. *If $x \in \mathcal{T}$, then $R(x)$ is a single geodesic.*

Proof. Since $x \in \mathcal{T}$, there exists $(w, x), (x, z) \in G \setminus \{x = y\}$: from the d_L-cyclical monotonicity and triangular inequality, it follows that

$$d_L(w, z) = d_L(w, x) + d_L(x, z),$$

so that $(w, z) \in G$ and $x \in \gamma_{(w,z)}$. Hence from the non branching assumption the set

$$R(x) = \bigcup_{y \in G(x)} \gamma_{[x,y]} \cup \bigcup_{z \in G^{-1}(x)} \gamma_{[z,x]}$$

is a single geodesic. □

Proposition 1.14. *The set $R \cap \mathcal{T} \times \mathcal{T}$ is an equivalence relation on $\mathcal{T}$. The set G is a partial order relation on $\mathcal{T}_e$.*

Proof. Using the definition of R, one has in $\mathcal{T}$:

1. $x \in \mathcal{T}$ implies that
$$\exists y \in G(x) \setminus \{x = y\},$$
so that from the definition of G it follows $(x, x) \in G$;

2. if $y \in R(x)$, $x, y \in \mathcal{T}$, then from Lemma 1.13 there exists $(w, z) \in G$ such that $x, y \in \gamma_{(w,z)}$. Hence $x \in R(y)$;

3. if $y \in R(x)$, $z \in R(y)$, $x, y, z \in \mathcal{T}$, then from Lemma 1.13 it follows again there exists $(w, z) \in G$ such that $x, y, z \in \gamma_{(w,z)}$. Hence $z \in R(x)$.

The second part follows similarly:

1. $x \in \mathcal{T}_e$ implies that
$$\exists (x, y) \in \big(G \setminus \{x = y\}\big) \cup \big(G^{-1} \setminus \{x = y\}\big),$$
so that in both cases $(x, x) \in G$;

2. as in Lemma 1.13, $(x, y), (y, z) \in G \setminus \{x = y\}$ implies by d_L-cyclical monotonicity that $(x, z) \in G$.

□

Remark 1.15. Note that $G \cup \{x = y\}$ is a partial order relation on X.

Definition 1.16. Define the multivalued *endpoint graphs* by:

$$a := \big\{(x, y) \in G^{-1} : G^{-1}(y) \setminus \{y\} = \emptyset\big\}, \tag{1.2.8a}$$
$$b := \big\{(x, y) \in G : G(y) \setminus \{y\} = \emptyset\big\}. \tag{1.2.8b}$$

We call $P_2(a)$ the set of *initial points* and $P_2(b)$ the set of *final points*.

Even if a, b are not in the analytic class, still they belong to the σ-algebra $\mathcal{A}$.

Proposition 1.17. *The following holds:*

1. *the sets*
$$a, b \subset X \times X, \quad a(A), b(A) \subset X,$$
belong to the $\mathcal{A}$-class if A analytic;

2. $a \cap b \cap \mathcal{T}_e \times X = \emptyset$;

3. $a(x)$, $b(x)$ *are singleton or empty when* $x \in \mathcal{T}$;

4. $a(\mathcal{T}) = a(\mathcal{T}_e)$, $b(\mathcal{T}) = b(\mathcal{T}_e)$;

5. $\mathcal{T}_e = \mathcal{T} \cup a(\mathcal{T}) \cup b(\mathcal{T})$, $\mathcal{T} \cap (a(\mathcal{T}) \cup b(\mathcal{T})) = \emptyset$.

Proof. Define
$$C := \Big\{(x, y, z) \in \mathcal{T}_e \times \mathcal{T}_e \times \mathcal{T}_e : y \in G(x), z \in G(y)\Big\} = (G \times X) \cap (X \times G) \cap \mathcal{T}_e \times \mathcal{T}_e \times \mathcal{T}_e,$$
that is clearly analytic. Then
$$b = \Big\{(x, y) \in G : y \in G(x), G(y) \setminus \{y\} = \emptyset\Big\} = G \setminus P_{1,2}(C \setminus X \times \{y = z\}),$$
$$b(A) = \Big\{y : y \in G(x), G(y) \setminus \{y\} = \emptyset, x \in A\Big\} = P_2(G \cap A \times X) \setminus P_2(C \setminus X \times \{y = z\}).$$
A similar computation holds for a:
$$a = G^{-1} \setminus P_{23}(C \setminus \{x = y\} \times X), \quad a(A) = P_1(G \cap X \times A) \setminus P_1(C \setminus \{x = y\} \times X)$$

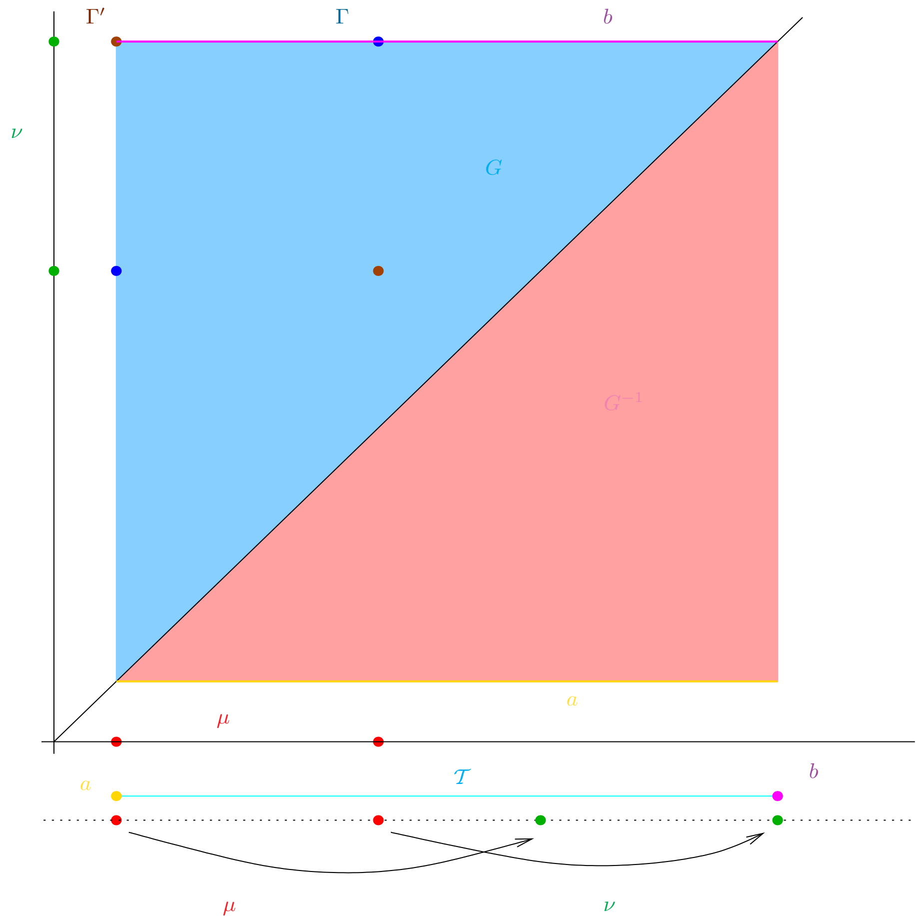

Figure 1.1: Construction of the sets Γ, Γ', G, G^{-1}, a, b in a 1-dimensional example with $d_L = |\cdot|$.

Hence $a, b \in \mathcal{A}(X \times X)$, $a(A), b(A) \in \mathcal{A}(X)$, being the intersection of an analytic set with a coanalytic one.

If $x \in \mathcal{T}$, then from Lemma 1.13 it follows that $a(x)$, $b(x)$ are empty or singletons and $a(x) \neq b(x)$. If $x \in \mathcal{T}_e \setminus \mathcal{T}$, then it follows that the geodesic $\gamma_{[w,z]}$, $(w, z) \in G$, to which x belongs cannot be prolonged in at least one direction: hence $x \in a(x) \cup b(x)$.

The other point follows easily. □

We finally show that we can assume that the μ-measure of final points and the ν-measure of the initial points are 0.

Lemma 1.18. *The sets $G \cap b(\mathcal{T}) \times X$, $G \cap X \times a(\mathcal{T})$ is a subset of the graph of the identity map.*

Proof. From the definition of b one has that

$$x \in b(\mathcal{T}) \quad \Longrightarrow \quad G(x) \setminus \{x\} = \emptyset,$$

A similar computation holds for a. □

Hence we conclude that

$$\pi(b(\mathcal{T}) \times X) = \pi(G \cap b(\mathcal{T}) \times X) = \pi(\{x = y\}),$$

and following (1.2.7) we can assume that

$$\mu(b(\mathcal{T})) = \nu(a(\mathcal{T})) = 0.$$

Remark 1.19. In the case considered in Remark 1.6, it is possible to obtain more regularity for the sets introduced so far. Recall that we are now assuming

(2') $d_L : X \times X \to [0, +\infty]$ l.s.c. distance,

(4') $d_L(x, y) \geq d(x, y)$,

(5') $\cup_{x \in K_1, y \in K_2} \gamma_{[x,y]}$ is d-compact if K_1, K_2 are d-compact, $d_{L \llcorner K_1 \times K_2}$ uniformly bounded.

The set Γ' is σ-compact: in fact, if one restrict to each Γ_n given by (1.2.1), then the set of cycles of order I is compact, and thus

$$\begin{aligned} \Gamma'_{n,\bar{I}} := \Big\{ (x, y) : \exists I \in \{0, \dots, \bar{I}\}, (w_i, z_i) \in \Gamma_n \text{ for } i = 0, \dots, I,\ z_I = y \\ w_{I+1} = w_0 = x,\ \sum_{i=0}^{I} d_L(w_{i+1}, z_i) - d_L(w_i, z_i) = 0 \Big\} \end{aligned}$$

is compact. Finally $\Gamma' = \cup_{n,I} \Gamma'_{n,I}$.

Moreover, $d_{L \llcorner \Gamma'_{n,I}}$ is continuous. If $(x_n, y_n) \to (x, y)$, then from the l.s.c. and

$$\sum_{i=0}^{I} d_L(w_{n,i+1}, z_{n,i}) = \sum_{i=0}^{I} d_L(w_{n,i}, z_{n,i}), \quad w_{n,I+1} = w_{n,0} = x_n,\ z_{n,I} = y_n,$$

it follows also that each $d_L(w_{n,i+1}, z_{n,i})$ is continuous.

Similarly the sets G, R, a, b are σ-compact: assumption (5') and the above computation in fact shows that

$$G_{n,I} := \Big\{ (x, y) : \exists (w, z) \in \Gamma'_{n,I}, d_L(w, x) + d_L(x, y) + d_L(y, z) = d_L(w, z) \Big\}$$

is compact. For a, b, one uses the fact that projection of σ-compact sets is σ-compact.

So if we are in the case of Remark 1.6, Γ, Γ', G, G^{-1}, a and b are σ-compact sets.

Remark 1.20. Many simplifications occur in the case the disintegration w.r.t. the partition $\{D_L(x)\}_{x\in X}$ is strongly consistent. Recall that $D_L(x) = \big\{y : d_L(x,y) < +\infty\big\}$. Let

$$\pi = \int_0^1 \pi_\alpha m(d\alpha), \quad \mu = \int_0^1 \mu_\alpha m(d\alpha),\ \nu = \int_0^1 \nu_\alpha m(d\alpha)$$

be strongly consistent disintegrations such that

$$\mu_\alpha(D_L(x_\alpha)) = \nu_\alpha(D_L(x_\alpha)) = 1, \quad \pi_\alpha \in \Pi(\mu_\alpha, \nu_\alpha).$$

We have used the fact that the partition $\{D_L(x) \times D_L(x)\}_{x\in X}$ has the crosswise structure, and then we can apply the results of [8].

1) Optimality of π_α. Since π is d_L-cyclically monotone, also the π_α are d_L-cyclically monotone: precisely they are concentrated on the sets

$$\Gamma_\alpha = \Gamma \cap D_L(x_\alpha) \times D_L(x_\alpha),$$

if Γ is d_L-cyclically monotone and $\pi(\Gamma) = 1$.

Using the fact that $(D_L(x_\alpha), d_L)$ is a metric space, then we can construct a potential $\varphi(x, x_\alpha)$ using the formula

$$\varphi(x, x_\alpha) = \inf\bigg\{\sum_{i=0}^{I} d_L(x_{i+1}, y_i) - d_L(x_i, y_i), (x_i, y_i) \in \Gamma_\alpha, x_{I+1} = x, (x_0, y_0) = (x_\alpha, x_\alpha)\bigg\}.$$

and since this is bounded on $(D_L(x_\alpha), d_L)$, we see that π_α and hence π are optimal.

2. Potential for π. Extend $\varphi(x, x_\alpha)$ to X by setting $\varphi(x, x_\alpha) = +\infty$ if $x \notin D_L(x_\alpha)$. If $\{(x_\alpha, x_\alpha)\}_{\alpha\in[0,1]}$ is a Borel section, then the function

$$\varphi(x) = \inf_\alpha\{\varphi(x, \alpha)\}$$

is easily seen to be analytic. This function is clearly a potential for π. In particular, it follows again from [8] that π is optimal if it is d_L-cyclically monotone.

3.Transport set. We can then define the set of oriented transport rays as the set

$$G = \Big\{(x,y) \in X \times X : \varphi(x) - \varphi(y) = d_L(x,y)\Big\}.$$

In general, this sets is larger than the one of Definition 1.9.

1.3 Partition of the transport set $\mathcal{T}$

In this section we use the continuity and local compactness of geodesics to show that the disintegration induced by R on $\mathcal{T}$ is strongly consistent. Using this fact, we can define an order preserving map g which maps our transport problem into a transport problem on $\mathcal{S} \times \mathbb{R}$, where $\mathcal{S}$ is a cross section of R.

Let $\{x_i\}_{i\in\mathbb{N}}$ be a dense sequence in (X, d).

Lemma 1.21. *The sets*

$$W_{ijk} := \Big\{x \in \mathcal{T} \cap \bar{B}_{2^{-j}}(x_i) : L(G(x)), L(G^{-1}(x)) \geq 2^{2-k}, L\big(R(x) \cap \bar{B}_{2^{1-j}}(x_i)\big) \leq 2^{-k}\Big\}$$

form a countable covering of $\mathcal{T}$ of class $\mathcal{A}$.

Proof. We first prove the measurability. We consider separately the conditions defining W_{ijk}.

Point 1. The set

$$A_{ij} := \mathcal{T} \cap \bar{B}_{2^{-j}}(x_i)$$

is clearly analytic.

Point 2. The set

$$B_k := \big\{x \in \mathcal{T} : L(G(x)) \geq 2^{2-k}\big\} = P_1\Big(G \cap \big\{d_L(x,y) \geq 2^{2-k}\big\}\Big)$$

is again analytic, being the projection of an analytic set. Similarly, the set

$$C_k := \big\{x \in \mathcal{T} : L(G^{-1}(x)) \geq 2^{2-k}\big\} = P_1\Big(G^{-1} \cap \big\{d_L(x,y) \geq 2^{2-k}\big\}\Big)$$

is again analytic.

Point 3. The set

$$\begin{aligned} D_{jk} &:= \big\{x \in \mathcal{T} : L\big(R(x) \cap \bar{B}_{2^{-j}}(x_i)\big) \leq 2^{-k}\big\} \\ &= \mathcal{T} \setminus P_1\Big(R \cap \big(\{(x,y) : d(x_i,y) \leq 2^{1-j}\} \cap \{d_L(x,y) > 2^{-k}\}\big)\Big) \end{aligned}$$

is in the $\mathcal{A}$-class, being the difference of two analytic sets.

We finally can write

$$W_{ijk} = A_{ij} \cap B_k \cap C_k \cap D_{jk},$$

and the fact that $\mathcal{A}$ is a σ-algebra proves that $W_{ijk} \in \mathcal{A}$.

To show that it is a covering, notice that for all $x \in \mathcal{T}$ it holds

$$\min\big\{L(G(x)), L(G^{-1}(x))\big\} \geq 2^{2-\bar{k}}$$

for some $\bar{k} \in \mathbb{N}$.

From the local compactness of geodesics, Condition 5. of page 19, it follows that if $\gamma^{-1}(\bar{B}_r(x))$ is compact, then the continuity of γ implies that $\gamma^{-1}(\bar{B}_{r'}(x))$ is also compact for all $r' \leq r$, and $\operatorname{diam}_{d_L}(\gamma \cap \bar{B}_{r'}(x)) \to 0$ and $r' \to 0$. In particular there exists $\bar{j} \in \mathbb{N}$ such that

$$L\big(R(x) \cap \bar{B}_{2^{1-\bar{j}}}(x)\big) \leq 2^{-\bar{k}},$$

with $\bar{k}$ the one chosen above.

Finally, one choose $x_{\bar{i}}$ such that $d(x, x_{\bar{i}}) < 2^{-1-\bar{j}}$, so that $x \in \bar{B}_{2^{-\bar{j}}}(x_{\bar{i}}) \subset \bar{B}_{2^{1-\bar{j}}}(x)$ and thus

$$L\big(R(x) \cap \bar{B}_{2^{-\bar{j}}}(x_{\bar{i}})\big) \leq 2^{-\bar{k}}.$$

□

Lemma 1.22. *There exist μ-negligible sets $N_{ijk} \subset W_{ijk}$ such that the family of sets*

$$\mathcal{T}_{ijk} = R^{-1}(W_{ijk} \setminus N_{ijk})$$

is a countable covering of $\mathcal{T} \setminus \cup_{ijk} N_{ijk}$ into saturated analytic sets.

Proof. First of all, since $W_{ijk} \in \mathcal{A}$, then there exists μ-negligible set $N_{ijk} \subset W_{ijk}$ such that $W_{ijk} \setminus N_{ijk} \in \mathcal{B}(X)$. Hence $\{W_{ijk} \setminus N_{ijk}\}_{i,j,k\in\mathbb{N}}$ is a countable covering of $\mathcal{T} \setminus \cup_{ijk} N_{ijk}$. It follows immediately that $\{\mathcal{T}_{ijk}\}_{i,j,k\in\mathbb{N}}$ satisfies the lemma. □

Remark 1.23. Observe that $\bar{B}_{2^{-j}}(x_i) \cap R(x)$ is compact for all $x \in \mathcal{T}_{ijk}$: in fact, during the proof of Lemma 1.21 we have already shown that $\gamma^{-1}(\bar{B}_{2^{-j}}(x_i))$ is compact.

From any analytic countable covering, we can find a countable partition into $\mathcal{A}$-class saturated sets by defining

$$\mathcal{Z}_{m,e} := \mathcal{T}_{i_m j_m k_m} \setminus \bigcup_{m'=1}^{m-1} \mathcal{T}_{i_{m'} j_{m'} k_{m'}}, \quad \mathcal{Z}_{0,e} := \mathcal{T}_e \setminus \bigcup_{m\in\mathbb{N}} \mathcal{Z}_{m,e}, \tag{1.3.1}$$

where

$$\mathbb{N} \ni m \mapsto (i_m, j_m, k_m) \in \mathbb{N}^3$$

is a bijective map. Intersecting the above sets with $\mathcal{T}$, we obtain the countable partition of $\mathcal{T}$ in $\mathcal{A}$-sets

$$\mathcal{Z}_m := \mathcal{Z}_{m,e} \cap \mathcal{T}, \quad m \in \mathbb{N}_0. \tag{1.3.2}$$

Now we use this partition to prove the strong consistency of the disintegration.

On $\mathcal{Z}_m$, $m > 0$, we define the closed values map

$$\mathcal{Z}_m \ni x \mapsto F(x) := R(x) \cap \bar{B}_{2^{-j_m}}(x_{i_m}) \in \mathcal{K}\big(\bar{B}_{2^{-j_m}}(x_{i_m})\big), \tag{1.3.3}$$

where $\mathcal{K}(\bar{B}_{2^{-j_m}}(x_{i_m}))$ is the space of compact subsets of $\bar{B}_{2^{-j_m}}(x_{i_m})$.

Proposition 1.24. *There exists a μ-measurable cross section $f : \mathcal{T} \to \mathcal{T}$ for the equivalence relation R.*

Proof. First we show that F is $\mathcal{A}$-measurable: for $\delta > 0$,

$$\begin{aligned} F^{-1}(B_\delta(y)) &= \Big\{ x \in \mathcal{Z}_m : R(x) \cap B_\delta(y) \cap \bar{B}_{2^{-j_m}}(x_{i_m}) \neq \emptyset \Big\} \\ &= \mathcal{Z}_m \cap P_1\Big(R \cap \big(X \times B_\delta(y) \cap \bar{B}_{2^{-j_m}}(x_{i_m}) \big) \Big). \end{aligned}$$

Being the intersection of two $\mathcal{A}$-class sets, $F^{-1}(B_\delta(y))$ is in $\mathcal{A}$.

By Corollary 5.7 there exists a $\mathcal{A}$-class section $f_m : \mathcal{Z}_m \to \bar{B}_{2^{-j_m}}(x_{i_m})$. The proposition follows by setting $f\llcorner_{\mathcal{Z}_m} = f_m$ on $\cup_m \mathcal{Z}_m$, and defining it arbitrarily on $\mathcal{T} \setminus \cup_m \mathcal{Z}_m$: the latter being negligible, f is μ-measurable. □

Up to a μ-negligible saturated set $\mathcal{T}_N$, we can assume it to have σ-compact range: just let $S \subset f(\mathcal{T})$ be a σ-compact set where $f_\sharp \mu\llcorner_{\mathcal{T}}$ is concentrated, and set

$$\mathcal{T}_S := R^{-1}(S) \cap \mathcal{T}, \quad \mathcal{T}_N := \mathcal{T} \setminus \mathcal{T}_S, \quad \mu(\mathcal{T}_N) = 0. \tag{1.3.4}$$

Having the $\mu\llcorner_{\mathcal{T}}$-measurable cross-section

$$\mathcal{S} := f(\mathcal{T}) = S \cup f(\mathcal{T}_N) = (\text{Borel}) \cup (f(\mu\text{-negligible})),$$

we can define the parametrization of $\mathcal{T}$ and $\mathcal{T}_e$ by geodesics.

Definition 1.25 (Ray map)**.** Define the *ray map* g by the formula

$$\begin{aligned} g := &\Big\{ (y,t,x) : y \in \mathcal{S}, t \in [0,+\infty), x \in G(y) \cap \{d_L(x,y) = t\} \Big\} \\ &\cup \Big\{ (y,t,x) : y \in \mathcal{S}, t \in (-\infty,0), x \in G^{-1}(y) \cap \{d_L(x,y) = -t\} \Big\} \\ = &\; g^+ \cup g^-. \end{aligned}$$

Proposition 1.26. *The following holds.*

1. *The restriction $g \cap S \times \mathbb{R} \times X$ is analytic.*
2. *The set g is the graph of a map with range $\mathcal{T}_e$.*
3. *$t \mapsto g(y,t)$ is a d_L 1-Lipschitz G-order preserving for $y \in \mathcal{T}$.*

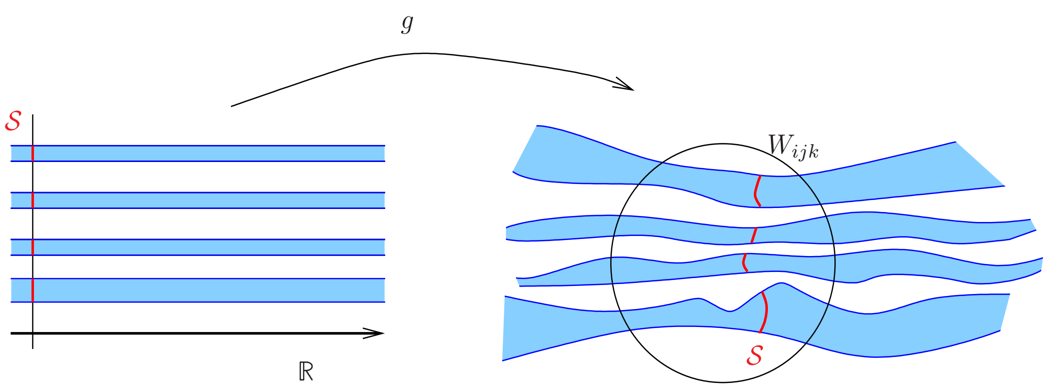

Figure 1.2: The ray map g.

4. *$(t, y) \mapsto g(y, t)$ is bijective on $\mathcal{T}$, and its inverse is*

$$x \mapsto g^{-1}(x) = \big(f(y), \pm d_L(x, f(y))\big)$$

where f is the quotient map of Proposition 1.24 and the positive/negative sign depends on $x \in G(f(y))/x \in G^{-1}(f(y))$.

Proof. For the first point just observe that

$$\begin{aligned} g^+ &= \Big\{(y,t,x) : y \in S, t \in \mathbb{R}^+, x \in G(y) \cap \{d_L(x,y) = t\}\Big\} \\ &= S \times R^+ \times X \cap \{(y,t,x) : (y,x) \in G\} \cap \{(y,t,x) : d_L(x,y) = t\} \in \Sigma^1_1. \end{aligned}$$

Similarly

$$g^- = \Big\{(y,t,x) : y \in S, t \in \mathbb{R}^-, x \in G^{-1}(y) \cap \{d_L(x,y) = -t\}\Big\} \in \Sigma^1_1.$$

Since $\mathcal{S} \subset \mathcal{T}$ and $R(y)$ is a subset of a single geodesic for $y \in \mathcal{S} \subset \mathcal{T}$, g is the graph of a map. Note that for any $x \in \mathcal{T}_e$ there exists $z \in \mathcal{T}$ such that $x \in R(z)$: hence $x \in R(f(z))$, and therefore the range of the map is the whole $\mathcal{T}_e$.

The third point is a direct consequence of the definition.The fourth point follows by substitution. □

We finally prove the following property of d_L-cyclically monotone transference plans.

Proposition 1.27. *For any π d_L-monotone there exists a d_L-cyclically monotone transference plan $\tilde{\pi}$ with the same cost of π such that it coincides with the identity on $\mu \wedge \nu$.*

We will use the disintegration technique exploited also in the next section. We observe that another proof can be the direct composition of the transference plan with itself, using the fact that the mass moves along geodesics and the disintegration makes the problem one dimensional.

Proof. We have already shown that we can take

$$\mu(P_2(b)) = \nu(P_2(a)) = 0,$$

so that $\mu \wedge \nu$ is concentrated on $\mathcal{T}_S$.

Step 1. On $\mathcal{T}$ we can use the Disintegration Theorem to write

$$\mu \llcorner_{\mathcal{T}} = \int_S \mu_y m(dy), \quad m = f_\sharp(\mu \llcorner_{\mathcal{T}}), \; \mu_y \in \mathcal{P}(R(y) \cap \mathcal{T}). \tag{1.3.5}$$

In fact, the existence of a Borel section is equivalent to the strong consistency of the disintegration. Since $\{R(y) \times X\}_{y \in \mathcal{T}}$ is also a partition on $\mathcal{T} \times X$, we can similarly write

$$\pi \llcorner_{\mathcal{T} \times X} = \int_S \pi_y m(dy), \quad \pi_y(R(y) \times R(y)) = 1.$$

We write moreover

$$\nu_y := (P_2)_\sharp (\pi \llcorner_{\mathcal{T} \times X}), \quad \tilde{\nu} := \int_S \nu_y m(dy) = \int_S (P_2)_\sharp \pi_y m(dy). \tag{1.3.6}$$

Clearly the rest of the mass starts from $a(\mathcal{T})$, so we have just to show how to rearrange the transference plan in $\mathcal{T}$ in order to obtain $\mu \perp \nu$. Using g, we can reduce the problem to a transport problem on $S \times \mathbb{R}$ with cost

$$c((y,t),(y',t')) = \begin{cases} |t - t'| & y = y' \\ +\infty & y \neq y' \end{cases}$$

By standard regularity argument, we can assume that $S \ni y \mapsto \pi_y \in \mathcal{P}(R(y) \times R(y))$ is σ-continuous, i.e. its graph is σ-compact.

Step 2. Using the fact that $(\mu, \nu) \mapsto \mu \wedge \nu$ is Borel w.r.t. the weak topology [8], we can assume that $S \ni y \mapsto \mu_y \wedge \nu_y \in \mathcal{P}(R(y))$ is σ-continuous, so that also the map

$$S \ni y \mapsto (\mu_y - \mu_y \wedge \nu_y, \nu_y - \mu_y \wedge \nu_y) \in \mathcal{P}(R(y)) \times \mathcal{P}(R(y))$$

is σ-continuous.

Step 3. Since in each $R(y)$ the problem is one dimensional, one can take the unique transference plan

$$\tilde{\pi}_y \in \Pi\big(\mu_y - \mu_y \wedge \nu_y, \nu_y - \mu_y \wedge \nu_y\big)$$

concentrated on a monotone set: clearly

$$\int d_L \tilde{\pi}_y = \int d_L \pi_y.$$

Step 4. If we define the left-continuous distribution functions

$$H(y,s) := \big(\mu_y - \mu_y \wedge \nu_y\big)(-\infty, s), \quad F(y,t) := \big(\nu_y - \mu_y \wedge \nu_y\big)(-\infty, t),$$

and

$$G(y,s,t) := \tilde{\pi}_y\big((-\infty,s) \times (-\infty,t)\big),$$

then the measure $\tilde{\pi}_y$ is uniquely determined by $G(y,s,t) = \min\{H(y,s), F(y,t)\}$.

The σ-continuity of $y \mapsto (\mu_y - \mu_y \wedge \nu_y, \nu_y - \mu_y \wedge \nu_y)$ yields that H, F are again σ-l.s.c., so that G is Borel, and finally $y \mapsto \tilde{\pi}_y$ is σ-continuous up to a $f_\sharp \mu$-negligible set.

Step 5. Define

$$\hat{\pi}_y := \tilde{\pi}_y + (\mathbb{I}, \mathbb{I})_\sharp (\mu_y \wedge \nu_y) \in \Pi(\mu_y, \nu_y).$$

The above steps show that $\hat{\pi}$ is m-measurable, and thus we can define the measure

$$\hat{\pi} := \pi \llcorner_{(\mathcal{T}_e \setminus \mathcal{T}) \times X} + \int \hat{\pi}_y m(dy).$$

It is routine to check that $\hat{\pi}$ has the required properties. $\square$

1.4 Regularity of the disintegration

This section is divided in two parts.

In the first one we consider the translation of Borel sets by the optimal geodesic flow, we introduce a first regularity assumption (Assumption 1) on the measure μ and we show that an immediate consequence is that the set of initial points is negligible. A second consequence is that the disintegration of μ w.r.t. R has continuous conditional probabilities.

In the second part we consider a stronger regularity assumption (Assumption 2) which gives that the conditional probabilities are absolutely continuous with respect to $\mathcal{H}^1$ along geodesics.

1.4.1 Evolution of Borel sets

Let $A \subset \mathcal{T}_e$ be an analytic set and define for $t \in \mathbb{R}$ the *t-evolution A_t of A* by

$$A_t := g\big(g^{-1}(A) + (0,t)\big). \tag{1.4.1}$$

Lemma 1.28. *The set $A_t \cap g(S \times \mathbb{R})$ is analytic, and A_t is μ-measurable for $t \geq 0$.*

Proof. Divide A into two parts:

$$A_S := A \cap g(S \times \mathbb{R}) \quad \text{and} \quad A_N := A \setminus A_S.$$

From Point (1) of Proposition 1.26 it follows that A_S is analytic. We consider the evolution of the two sets separately.

Again by Point (1) of Proposition 1.26, the set $(A_S)_t$ is analytic, hence universally measurable for all $t \in \mathbb{R}$.

Since $\mathcal{T}_N$ is μ-negligible (see (1.3.4)), it follows that $(A_N)_t$ is μ-negligible for all $t > 0$, and by the assumptions it is clearly measurable for $t = 0$. □

We can show that $t \mapsto \mu(A_t)$ is measurable.

Lemma 1.29. *Let A be analytic. The function $t \mapsto \mu(A_t)$ is Souslin for $t \geq 0$. If $A \subset g(S \times \mathbb{R})$, then $t \mapsto \mu(A_t)$ is Souslin for $t \in \mathbb{R}$.*

Proof. As before, we split the A into the sets

$$A_S := A \cap g(S \times \mathbb{R}) \quad \text{and} \quad A_N := A \setminus A_S.$$

The function

$$t \mapsto \mu(A_{N,t}) = \begin{cases} \mu(A_N) & t = 0 \\ 0 & t > 0 \end{cases}$$

is clearly Borel. Observe that since $\mathcal{T}_N \subset \mathcal{T}$ and the μ-measure of final points is 0, the value of $\mu(A_{N,t})$ is known only for $t > 0$.

Since A_S is analytic, then $g^{-1}(A_S)$ is analytic, and the set

$$\tilde{A}_S := \big\{(y,\tau,t) : (y,\tau - t) \in g^{-1}(A_S)\big\}$$

is easily seen to be again analytic. Define the analytic set $\hat{A}_S \subset X \times \mathbb{R}$ by

$$\hat{A}_S := (g, \mathbb{I})(\tilde{A}_S).$$

Clearly $(A_S)_t = \hat{A}_S(t)$. We now show in two steps that the function $t \mapsto \mu((A_S)_t)$ is analytic.

Step 1. Define the closed set in $\mathcal{P}(X \times [0,1])$

$$\Pi(\mu) := \big\{\pi \in \mathcal{P}(X \times [0,1]) : (P_1)_\sharp(\pi) = \mu\big\}$$

and let $B \subset X \times \mathbb{R} \times [0,1]$ a Borel set such that $P_{12}(B) = \hat{A}_S$.

Consider the function

$$\mathbb{R} \times \Pi(\mu) \ni (t, \pi) \mapsto \pi(B(t)).$$

A slight modification of Lemma 4.12 in [8] shows that this function is Borel.

Step 2. Since supremum of Borel function are Souslin, pag. 134 of [25], the proof is concluded once we show that

$$\mu((A_S)_t) = \mu(\hat{A}_S(t)) = \sup_{\pi \in \Pi(\mu)} \pi(B(t)).$$

From the Disintegration Theorem, for all $\pi \in \Pi(\mu)$ we have

$$\pi(B(t)) = \int \pi_x(B(t))\mu(dx) \leq \int_{P_1(B(t))} \mu(dx) = \mu(\hat{A}_S(t)).$$

On the other hand from Theorem 5.5, there exists an $\mathcal{A}$-measurable section $u : \hat{A}_S(t) \to B(t)$. Clearly for $\pi_u = (\mathbb{I}, u)_\sharp(\mu)$ it holds $\pi_u(B(t)) = \mu(\hat{A}_S(t))$. □

The next assumption is the first fundamental assumption of the chapter.

Assumption 1 (Non-degeneracy assumption)**.** For all Borel sets A such that $\mu(A) > 0$ the set $\{t \in \mathbb{R}^+ : \mu(A_t) > 0\}$ has cardinality $> \aleph_0$.

By inner regularity, it is clearly enough to verifies Assumption 1 only for compact sets. Note that since for analytic set Cantor Hypothesis holds true, Theorem 4.3.5, pag. 142 of [25] , Assumption 1 implies that the cardinality of $\{t \in \mathbb{R}^+ : \mu(A_t) > 0\}$ is $\mathfrak{c}$.

An immediate consequence of the Assumption 1 is that the measure μ is concentrated on $\mathcal{T}$.

Lemma 1.30. *If μ satisfies Assumption 1 then*

$$\mu(\mathcal{T}_e \setminus \mathcal{T}) = 0.$$

Proof. If $A \subset a(X)$, then $A_t \cap A_s = \emptyset$ for $0 \leq s < t$. Hence

$$\sharp\big\{t \in \mathbb{R}^+ : \mu(A_t) > 0\big\} \leq \aleph_0,$$

because of the boundedness of μ. This contradicts the assumptions. □

Once we know that $\mu(\mathcal{T}) = 1$, we can use the Disintegration Theorem 5.3 to write

$$\mu = \int_S \mu_y m(dy), \quad m = f_\sharp \mu, \ \mu_y \in \mathcal{P}(R(y)). \tag{1.4.2}$$

The disintegration is strongly consistent since the quotient map $f : \mathcal{T} \to \mathcal{T}$ is μ-measurable and $(\mathcal{T}, \mathcal{B}(\mathcal{T}))$ is countably generated.

The second consequence of Assumption 1 is that μ_y is continuous, i.e. $\mu_y(\{x\}) = 0$ for all $x \in X$.

Proposition 1.31. *The conditional probabilities μ_y are continuous for m-a.e. $y \in S$.*

Proof. From the regularity of the disintegration and the fact that $m(S) = 1$, we can assume that the map $y \mapsto \mu_y$ is weakly continuous on a compact set $K \subset S$ of comeasure $< \epsilon$ such that $L(R(y)) > \epsilon$ for all $y \in K$. It is enough to prove the proposition on K.

Step 1. From the continuity of $K \ni y \mapsto \mu_y \in \mathcal{P}(X)$ w.r.t. the weak topology, it follows that the map

$$y \mapsto A(y) := \big\{x \in R(y) : \mu_y(\{x\}) > 0\big\} = \cup_n \big\{x \in R(y) : \mu_y(\{x\}) \geq 2^{-n}\big\}$$

is σ-closed: in fact, if $(y_m, x_m) \to (y, x)$ and $\mu_{y_m}(\{x_m\}) \geq 2^{-n}$, then $\mu_y(\{x\}) \geq 2^{-n}$ by u.s.c. on compact sets.

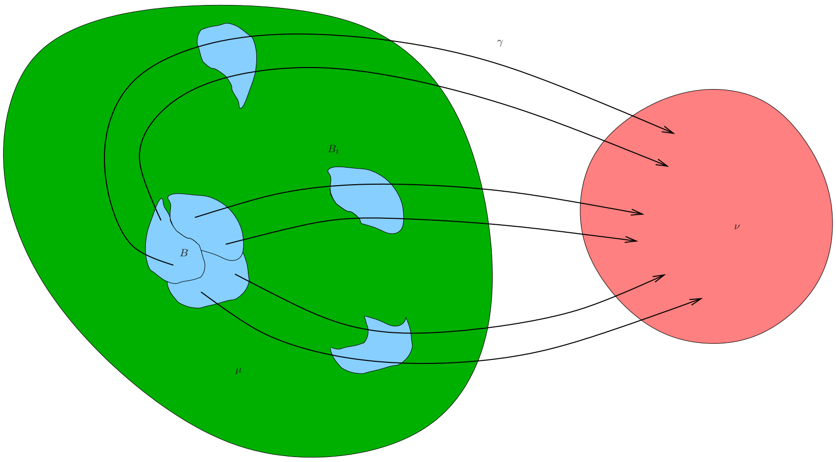

Figure 1.3: The evolution of a set B through the optimal flow.

Hence it is Borel, and by Lusin Theorem (Theorem 5.8.11 of [25]) it is the countable union of Borel graphs: setting in case $c_i(y) = 0$, we can consider them as Borel functions on S and order them w.r.t. G,

$$\mu_{y,\text{atomic}} = \sum_{i \in \mathbb{Z}} c_i(y) \delta_{x_i(y)}, \quad x_{i+1}(y) \in G(x_i(y)),\ i \in \mathbb{Z}.$$

Step 2. Define the sets

$$S_{ij}(t) := \Big\{ y \in K : x_i(y) = g\big(g^{-1}(x_j(y)) + t\big) \Big\} \cap \mathcal{T}.$$

Since $K \subset S$, to define S_{ij} we are using the graph $g \cap S \times \mathbb{R} \times \mathcal{T}$, which is analytic: hence $S_{ij} \in \Sigma^1_1$.

For $A_j := \{x_j(y), y \in K\}$ and $t \in \mathbb{R}^+$ we have that

$$\begin{aligned} \mu((A_j)_t) &= \int_K \mu_y((A_j)_t) m(dy) = \int_K \mu_{y,\text{atomic}}((A_j)_t) m(dy) \\ &= \sum_{i \in \mathbb{Z}} \int_K c_i(y) \delta_{x_i(y)} \big(g(g^{-1}(x_j(y)) + t)\big) m(dy) = \sum_{i \in \mathbb{Z}} \int_{S_{ij}(t)} c_i(y) m(dy). \end{aligned}$$

We have used the fact that $A_j \cap R(y)$ is a singleton.

Step 3. For fixed $i, j \in \mathbb{N}$, again from the fact that $A_j \cap R(y)$ is a singleton

$$S_{ij}(t) \cap S_{ij}(t') = \begin{cases} S_{ij}(t) & t = t' \\ \emptyset & t \neq t' \end{cases}$$

so that

$$\sharp \big\{ t : m(S_{ij}(t)) > 0 \big\} \leq \aleph_0.$$

Finally

$$\mu((A_j)_t) > 0 \quad \Longrightarrow \quad t \in \bigcup_i \big\{ t : m(S_{ij}(t)) > 0 \big\},$$

whose cardinality is $\leq \aleph_0$, contradicting Assumption 1. $\square$

1.4.2 Absolute continuity

We next assume a stronger regularity assumption.

Assumption 2 (Absolute continuity assumption)**.** For every Borel set $A \subset \mathcal{T}_e$

$$\mu(A) > 0 \quad \Longrightarrow \quad \int_0^{+\infty} \mu(A_t) dt > 0.$$

Again by inner regularity, Assumption 2 can be verified only for compact sets. Note that the condition is meaningful by Lemma 1.29. Observe moreover that Assumption 2 implies Assumption 1, so that in the following we will restrict the map g to the set $g^{-1}(\mathcal{T})$, where it is analytic. Moreover, we can consider shift $t \mapsto A_t$ for $t \in \mathbb{R}$, because of Lemma 1.29.

Remark 1.32. An equivalent form of the Assumption 2 is the following:

$$\mu(A) > 0 \quad \Longrightarrow \quad \int_{t,s \geq 0} \mu(A_t \cap A_s) dt ds > 0.$$

In fact, due to $\mu(X) = 1$, in the set $I_n := \{t : \mu(A_t) > 2^{-n}\}$ the set $\{s \in I_n : \mu(A_s \cap A_t) = 0, t \in I_n\}$ has cardinality at most 2^{-n}. Hence, since for some n $\mathcal{L}^1(I_n) > 0$ by Assumption 2, it follows that

$$\mathcal{L}^2(I_n \times I_n) = \big(\mathcal{L}^1(I_n)\big)^2 > 0.$$

The opposite implication is a consequence of Fubini theorem.

The next results show regularity of the Radon-Nikodym derivative of μ_y w.r.t. $(\mathcal{H}^1_L)\llcorner_{f^{-1}(y)}$, where $\mathcal{H}^1_L$ is the 1-dimensional Hausdorff measure w.r.t. the d_L-distance. Note that along d_L 1-Lipschitz geodesics, $\mathcal{H}^1_L$ is equivalence to $g(y,\cdot)_\sharp \mathcal{L}^1$: in the following we will use both notations.

Lemma 1.33. *Let μ be a Radon measure and*

$$\mu_y = r(y,\cdot) g(y,\cdot)_\sharp \mathcal{L}^1 + \omega_y, \quad \omega_y \perp g(y,\cdot)_\sharp \mathcal{L}^1$$

be the Radon-Nikodym decomposition of μ_y w.r.t. $g(y,\cdot)_\sharp \mathcal{L}^1$. Then there exists a Borel set $C \subset X$ such that

$$\mathcal{L}^1\big(g^{-1}(C) \cap (\{y\} \times \mathbb{R})\big) = 0$$

and $\omega_y = \mu_y \llcorner_C$ for m-a.e. $y \in [0,1]$.

Proof. Consider the measure

$$\lambda = g_\sharp(m \otimes \mathcal{L}^1),$$

and compute the Radon-Nikodym decomposition

$$\mu = \frac{D\mu}{D\lambda} \lambda + \omega.$$

Then there exists a Borel set C such that $\omega = \mu \llcorner_C$ and $\lambda(C) = 0$. The set C proves the Lemma. Indeed $C = \cup_{y \in [0,1]} C_y$ where $C_y = C \cap f^{-1}(y)$ is such that $\mu_y \llcorner_{C_y} = \omega_y$ and $g(y,\cdot)_\sharp \mathcal{L}^1(C_y) = 0$ for m-a.e. $y \in [0,1]$. □

Theorem 1.34. *If μ satisfies Assumption 2, then for m-a.e. $y \in [0,1]$ the conditional probabilities μ_y are absolutely continuous w.r.t. $g(y,\cdot)_\sharp \mathcal{L}^1$.*

The proof is based on the following simple observation.

Let η be a Radon measure on $\mathbb{R}$. Suppose that for all $A \subset \mathbb{R}$ Borel with $\eta(A) > 0$ it holds

$$\int_{\mathbb{R}^+} \eta(A + t) dt = \eta \otimes \mathcal{L}^1\big(\{(x,t) : t \geq 0, x - t \in A\}\big) > 0.$$

Then $\eta \ll \mathcal{L}^1$.

Proof. The proof will use Lemma 1.33: take C the set constructed in Lemma 1.33 and suppose by contradiction that

$$\mu(C) > 0 \quad \text{and} \quad m \otimes \mathcal{L}^1(g^{-1}(C)) = 0.$$

In particular, for all $t \in \mathbb{R}$ it follows that

$$m \otimes \mathcal{L}^1(g^{-1}(C_t)) = m \otimes \mathcal{L}^1(g^{-1}(C) + (0,t)) = 0.$$

By Fubini-Tonelli Theorem

$$\begin{aligned}
0 < \int_{\mathbb{R}^+} \mu(C_t)dt &= \int_{\mathbb{R}^+} \bigg(\int_{g^{-1}(C_t)} (g^{-1})_\sharp \mu(dyd\tau) \bigg) dt \\
&= \big((g^{-1})_\sharp \mu \otimes \mathcal{L}^1\big) \Big(\Big\{ (y,\tau,t) : (y,\tau) \in g^{-1}(\mathcal{T}), (y, \tau - t) \in g^{-1}(C) \Big\} \Big) \\
&\leq \int_{S\times\mathbb{R}} \mathcal{L}^1\big(\{\tau - g^{-1}(C \cap f^{-1}(y))\}\big)(g^{-1})_\sharp \mu(dyd\tau) \\
&= \int_{S\times\mathbb{R}} \mathcal{L}^1\big(g^{-1}(C \cap f^{-1}(y))\big)(g^{-1})_\sharp \mu(dyd\tau) \\
&= \int_{S} \mathcal{L}^1\big(g^{-1}(C \cap f^{-1}(y))\big) m(dy) = 0.
\end{aligned}$$

That gives a contradiction. □

Now we will study the regularity of the map $t \mapsto \mu(A_t)$ under Assumption 2. We will use the following notation:

$$\mu(A) = \int_S \mu_y(A) m(dy) = \int_S \bigg(\int_{g(y,\cdot)^{-1}(A)} r(y,\tau)d\tau \bigg) m(dy) = g_\sharp(rm \otimes \mathcal{L}^1).$$

Proposition 1.35. *μ satisfies Assumption 2 if and only if for all A Borel $t \mapsto \mu(A_t)$ is continuous. Moreover if A is geodesically convex then $\mu(A_t)$ is absolutely continuous.*

Proof. It is enough to prove the continuity for $t = 0$. Since

$$\mu(A_t) = \int_S \bigg(\int_{g(y,\cdot)^{-1}(A_t)} r(y,\tau)d\tau \bigg) m(dy),$$

its continuity is a direct consequence of Lebesgue dominated convergence theorem applied to the function:

$$t \mapsto \mu_y(A_t) = \int_{g(y,\cdot)^{-1}(A_t)} r(y,\tau)d\tau.$$

Suppose now A geodesically convex. Each $g(y,\cdot)^{-1}(A)$ is an interval $(\alpha(y), \omega(y))$, so that the map

$$t \mapsto \int_{g(y,\cdot)^{-1}(A_t)} r(y,\tau)d\tau$$

is absolutely continuous with derivative

$$h(y,t) = r(y, \omega(y) + t) - r(y, \alpha(y) + t).$$

Since $h(y,t) \in L^1(m \otimes \mathcal{L}^1)$ the result follows by a standard computation. □

1.5 Solution to the Monge problem

In this section we show that Theorem 1.34 allows to construct an optimal map T. We recall the one dimensional result for the Monge problem [31].

Theorem 1.36. *Let μ, ν be probability measures on $\mathbb{R}$, μ continuous, and let*

$$H(s) := \mu((-\infty, s)), \quad F(t) := \nu((-\infty, t)),$$

be the left-continuous distribution functions of μ and ν respectively. Then the following holds.

1. *The non decreasing function $T : \mathbb{R} \to \mathbb{R} \cup [-\infty, +\infty)$ defined by*

$$T(s) := \sup\big\{t \in \mathbb{R} : F(t) \leq H(s)\big\}$$

maps μ to ν. Moreover any other non decreasing map T' such that $T'_\sharp\mu = \nu$ coincides with T on the support of μ up to a countable set.

2. *If $\phi : [0, +\infty] \to \mathbb{R}$ is non decreasing and convex, then T is an optimal transport relative to the cost $c(s, t) = \phi(|s - t|)$. Moreover T is the unique optimal transference map if ϕ is strictly convex.*

Assume that μ satisfies Assumption 1. Then we can disintegrate μ and π respect to the ray equivalence relation R and $R \times X$ as in (1.4.2),

$$\mu = \int \mu_y m(dy), \ \pi = \int \pi_y m(dy), \quad \mu_y \text{ continuous}, \ (P_1)_\sharp \pi_y = \mu_y. \tag{1.5.1}$$

We write moreover

$$\nu = \int \nu_y m(dy) = \int (P_2)_\sharp \pi_y m(dy). \tag{1.5.2}$$

Note that $\pi_y \in \Pi(\mu_y, \nu_y)$ is d_L-cyclically monotone (and hence optimal, because $R(y)$ is one dimensional) for m-a.e. y. If $\nu(\mathcal{T}) = 1$, then (1.5.2) is the disintegration of ν w.r.t. R.

Theorem 1.37. *Let $\pi \in \Pi(\mu, \nu)$ be a d_L-cyclically monotone transference plan, and assume that Assumption 1 holds. Then there exists a Borel map $T : X \to X$ with the same transport cost as π.*

Proof. By means of the map g^{-1}, we reduce to a transport problem on $S \times \mathbb{R}$, with cost

$$c((y, s), (y', t)) = \begin{cases} |t - s| & y = y' \\ +\infty & y \neq y' \end{cases}$$

It is enough to prove the theorem in this setting under the following assumptions: S compact and $S \ni y \mapsto (\mu_y, \nu_y)$ weakly continuous. We consider here the probabilities μ_y, ν_y on $\mathbb{R}$.

Step 1. From the weak continuity of the map $y \mapsto (\mu_y, \nu_y)$, it follows that the maps

$$(y, t) \mapsto H(y, t) := \mu_y((-\infty, t)), \quad (y, t) \mapsto F(y, t) := \nu_y((-\infty, t))$$

are easily seen to be l.s.c.. Both are clearly increasing in t. Note also that H is continuous in t.

Step 2. The map T defined as Theorem 1.36 by

$$T(y, s) := \Big(y, \sup\big\{t : F(y, t) \leq H(y, s)\big\}\Big)$$

is Borel. In fact, for A Borel,

$$T^{-1}(A \times [t, +\infty)) = \big\{(y, s) : y \in A, H(y, s) \geq F(y, t)\big\} \in \mathcal{B}(S \times \mathbb{R}).$$

Step 3. Note that π_y and $T(y, \cdot)$ are both optimal for the transport problem between μ_y and ν_y with cost d_L restricted to $R(y)$. Indeed d_L restricted to $R(y) \times R(y)$ is finite. Therefore π_y and $T(y, \cdot)$ have the same cost. ☐

Remark 1.38. By the definition of the set G, it follows that along each geodesic

$$\mu_y(g(y,(-\infty,t))) \geq \nu_y(g(y,(-\infty,t))),$$

because in the opposite case G is not d_L-cyclically monotone.

Therefore $T(s) \geq s$, and $c((y,s),T(y,s)) = P_2(T(y,s)) - s$. Hence

$$\begin{aligned}\int d_L\pi &= \int d_L(x,T(x))\mu(dx) \\ &= \int_{S\times\mathbb{R}} s\big(g(y,\cdot)_\sharp^{-1}(\nu_y-\mu_y)\big)(ds)m(dy) = \int P_2(g^{-1}(x))(\nu-\mu)(dx). \qquad (1.5.3)\end{aligned}$$

1.6 Dynamic interpretation

In this section we show how the regularity of the disintegration yields a correct definition of the current $\dot{g}$ representing the flow along the geodesics of an optimal transference plan. This allows to solve the PDE

$$\partial U = \mu - \nu$$

in the sense of currents in metric spaces. In particular, under additional regularity assumptions, one can prove that the boundary $\partial\dot{g}$ is well defined and satisfies an ODE along geodesics. This gives a dynamic interpretation to the transport problem.

The setting here is slightly different from the previous sections:

1. $d(x,y) \leq d_L(x,y)$;
2. there exists a probability measure η, such that it (or more precisely $\eta\llcorner_{\mathcal{T}_e}$) satisfies Assumption 2 along the transport rays of the transportation problem with marginals μ, ν;
3. $\mu \ll \eta$, so that also μ satisfies Assumption 2.

In particular, $\mathrm{Lip}(X) \subset \mathrm{Lip}_{d_L}(X)$.

The main reference for this chapter is [3].

1.6.1 Definition of $\dot{g}$

For any Lipschitz function $\omega : X \to \mathbb{R}$ we can define the derivative $\partial_t\omega$ along the geodesic $g(t,y)$ for a.e. $t \in \mathbb{R}$,

$$\partial_t\omega(g(y,t)) := \frac{d}{dt}\omega(g(t,y)).$$

Using the disintegration formula

$$\eta\llcorner_{\mathcal{T}} = \int (g(y,\cdot))_\sharp(q(y,\cdot)\mathcal{L}^1)m(dy) = g_\sharp(qm\otimes\mathcal{L}^1)$$

for some $q \in L^1(m\otimes\mathcal{L}^1)$ (Theorem 1.34), we can define the measure $\partial_t\omega\eta$ as

$$\int \phi(x)(\partial_t\omega\eta)(dx) := \int_S\int_{\mathbb{R}} \phi(g(y,t))\partial_t\omega(g(y,t))q(y,t)dtm(dy).$$

where $\phi \in C_b(X,\mathbb{R})$.

Definition 1.39. We define the *flow* $\dot{g}$ as the current

$$\langle\dot{g},(h,\omega)\rangle = \int_{S\times\mathbb{R}} h(g(y,t))\partial_t\omega(g(y,t))q(y,t)dtm(dy)$$

where h, ω are Lipschitz functions of (X,d) with h bounded.

It is fairly easy to see that $\dot g$ is a current: in fact,

1. $\dot g$ has finite mass, namely
$$\big|\langle \dot g, (h,\omega)\rangle\big| \leq \mathrm{Lip}(\omega) \int h\eta;$$
2. $\dot g$ is linear in h, ω;
3. if $\omega_n \to \omega$ pointwise in X with uniformly bounded Lipschitz constant, then by Lebesgue Dominated Convergence Theorem if follows that
$$\lim_{n\to+\infty} \langle \dot g, (h_n,\omega_n)\rangle = \langle \dot g, (h,\omega)\rangle;$$
4. $\langle \dot g, (h,\omega)\rangle = 0$ if ω is constant in $\{h \neq 0\}$.

In general, $\dot g$ is only a current, with boundary $\partial \dot g$ defined by the duality formula
$$\langle \partial \dot g, \omega\rangle = \langle \dot g, (1,\omega)\rangle. \tag{1.6.1}$$
Under additional assumptions, the current $\dot g$ is a normal current, i.e. $\partial \dot g$ is also a scalar current, in particular it is a bounded measure on (X,d).

Lemma 1.40. *Assume that $q(y,\cdot) : \mathbb{R} \to \mathbb{R}$ belongs to $\mathrm{BV}(\mathbb{R})$ for m-a.e. y and*
$$\sigma_y := -\frac{d}{dt} q(y,t), \qquad \int_S |\sigma_y(\mathbb{R})| m(dy) = \int_S \mathrm{Tot.Var.}(q(y,\cdot)) m(dy) < +\infty.$$
Then $\dot g$ is a normal current and its boundary is given by
$$\langle \partial \dot g, \omega\rangle = \int_S \int_{\mathbb{R}} \omega(g(y,t)) \sigma_y(dt) m(dy).$$

Note that in the above formula we cannot restrict σ_y to $g^{-1}(\mathcal{T})$: in fact, in general
$$\int_S (g(y,\cdot)_\sharp \sigma_y)(\mathcal{T}_e \setminus \mathcal{T}) m(dy) > 0.$$

Proof. First of all, by using the formula $q(y,t) = \sigma_y((t,+\infty))$, it follows that σ_y is m-measurable, i.e. for all $\phi \in C_b(X,\mathbb{R})$ the integral
$$\int \bigg(\int \phi(g(y,t)) \sigma_y(dt) \bigg) m(dy)$$
is meaningful and then
$$\int \big(g(y,\cdot)_\sharp \sigma_y\big) m(dy)$$
is a finite measure on (X,d).

A direct computation yields
$$\langle \partial \dot g, \omega\rangle = \langle \dot g, (1,\omega)\rangle = \int_S \int_{\mathbb{R}} \partial_t \omega(g(t,y)) \sigma_y((t,+\infty)) dt m(dy) = \int_S \int_{\mathbb{R}} \omega(g(t,y)) \sigma_y(dt) m(dy).$$
□

Remark 1.41. In many cases the measure $\int (g(y,\cdot)_\sharp \sigma_y) \llcorner_{\mathcal{T}} m(dy)$ is absolutely continuous w.r.t. η, i.e. for m-a.e. y
$$\sigma_y \llcorner_{\mathcal{T}} = h(g(t,y)) q(y,t) \mathcal{L}^1.$$
for some $h \in L^1(\eta)$. In that case we obtain that
$$\begin{aligned}\langle \partial \dot g, \omega\rangle = & \int \omega(b(y)) \sigma_y\big(P_2(\{g^{-1}(b(y))\})\big) m(dy) \\ & - \int \omega(a(y)) \sigma_y\big(P_2(\{g^{-1}(a(y))\})\big) m(dy) + \int \omega(x) h(x) \eta(dx).\end{aligned}$$

1.6.2 Transport equation

We now consider the problem $\partial U = \mu - \nu$ in the sense of currents:

$$\langle U, (1,\omega) \rangle = \langle \mu - \nu, \omega \rangle = \int \omega(x)(\mu - \nu)(dx).$$

Using the disintegration formula and (1.5.1), (1.5.2) we can write

$$\langle U, (1,\omega) \rangle = \int_S \bigg\{ \int_{\mathbb{R}} \omega(g(y,t))(g^{-1}(y,\cdot)_\sharp \mu_y)(dt) - \int_{\mathbb{R}} \omega(g(y,t))(g^{-1}(y,\cdot)_\sharp \nu_y)(dt) \bigg\} m(dy).$$

By integrating by parts we obtain

$$\begin{aligned} \int_{\mathbb{R}} \omega(g(y,t))(g^{-1}(y,\cdot)_\sharp \mu_y)(dt) &= -\int_{\mathbb{R}} \mu_y(g(y,(-\infty,t)))\partial_t \omega(g(y,t))dt \\ &= -\int_{\mathbb{R}} H(y,t)\partial_t \omega(g(y,t))dt, \end{aligned}$$

$$\begin{aligned} \int_{\mathbb{R}} \omega(g(y,t))(g^{-1}(y,\cdot)_\sharp \nu_y)(dt) &= -\int_{\mathbb{R}} \nu_y(g(y,(-\infty,t)))\partial_t \omega(g(y,t))dt \\ &= -\int_{\mathbb{R}} F(y,t)\partial_t \omega(g(y,t))dt. \end{aligned}$$

Observe that the map

$$S \times \mathbb{R} \ni (y,t) \mapsto F(y,t) - H(y,t) \in \mathbb{R}$$

is in $L^1(m \otimes \mathcal{L}^1)$ if the transport cost $\mathcal{I}(\pi)$ is finite: in fact, using the fact that $F(y,t) \leq H(y,t)$ and integrating by parts,

$$\int_{\mathbb{R}} H(y,t) - F(y,t)dt = \int_{\mathbb{R}} (g^{-1}(y,\cdot)_\sharp \mu_y - g^{-1}(y,\cdot)_\sharp \nu_y)(-\infty,t)(dt) = \int_{\mathbb{R}^2} (t-s)\tilde{\pi}_y(ds,dt), \tag{1.6.2}$$

where $\tilde{\pi}_y$ is the monotone rearrangement.

We deduce the following proposition.

Proposition 1.42. *Under Assumption 1, a solution to $\partial U = \mu - \nu$ is given by the current U defined as*

$$\langle U, (h,\omega) \rangle = \int_S \bigg(\int_{\mathbb{R}} (F(y,t) - H(y,t))h(g(y,t))\partial_t \omega(g(y,t))dt \bigg) m(dy).$$

In general, the solution is not unique: just add a boundary free current to our solution.

Some further assumptions allow to represent our solution U as the product of a scalar ρ with the current $\dot{g}$.

Proposition 1.43. *Assume that $q(y,t) > 0$ whenever $H(y,t) - F(y,t) > 0$. Then $R = \rho \dot{g}$, where*

$$\rho(g(y,t)) = \frac{F(y,t) - H(y,t)}{q(y,t)}.$$

Proof. It is enough to observe that

$$\begin{aligned} \int_{S \times \mathbb{R}} F(y,t) - H(y,t)dtm(dy) &= \int_{S \times \mathbb{R}} \frac{F(y,t) - H(y,t)}{q(y,t)} q(y,t)dtm(dy) \\ &= \int_{S \times \mathbb{R}} \rho(g(y,t))q(y,t)dtm(dy) = \int_X \rho(x)\eta(dx), \end{aligned}$$

and from (1.6.2) we conclude that $\rho \in L^1(\eta)$. $\square$

Corollary 1.44. *If $q(y,t) \neq 0$ for $m \otimes \mathcal{L}^1$-a.e. $(y,t) \in g^{-1}(\mathcal{T})$, then there exists a scalar function ρ such that $\partial(\rho \dot{g}) = \mu - \nu$.*

1.7 Stability of the non degeneracy condition

In this section we prove a general approximation theorem, which will be then applied to the Measure-Gromov-Hausdorff (MGH) convergence: if a uniform estimate holds for the disintegration in the approximating spaces, we deduce the regularity of the disintegration also in the limit.

1.7.1 A general stability result

We consider the following setting:

1. μ_n is a sequence of measure converging to μ weakly;
2. there exists functions $g_n : S_n \times \mathbb{R} \to X$, $S_n \subset X$ Borel, and measures $r_n m_n \otimes \mathcal{L}^1 \in \mathcal{P}(S_n \times \mathbb{R})$ such that

$$\mu_n = (g_n)_\sharp \big(r_n m_n \otimes \mathcal{L}^1 \big). \tag{1.7.1}$$

The following is the basic tool for our stability result.

Proposition 1.45. *Let Y be a Polish space, $\{\xi_n\}_{n\in\mathbb{N}} \subset \mathcal{P}(Y)$ such that $\xi_n \rightharpoonup \xi$. Consider $\{r_n\}_{n\in\mathbb{N}}$, $r_n \geq 0$, such that $r_n \in L^1(\xi_n)$, $r_n \xi_n \rightharpoonup \zeta$ and the following equintegrability condition holds:*

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \bigg(\forall A \in \mathcal{B}, \xi_n(A) < \delta \quad \Longrightarrow \quad \int_A r_n \xi_n < \varepsilon \bigg).$$

Then there exists $r \in L^1(\xi)$ such that $\zeta = r\xi$.

Proof. We will show that $\zeta(B) = 0$ for all B such that $\xi(B) = 0$. Clearly by inner and outer regularity, it is enough to prove the following statement:

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \bigg(\phi \in C_b(Y), \phi \geq 0, \ \int \phi \xi < \delta \quad \Longrightarrow \int \phi \zeta < \varepsilon \bigg).$$

Fix $\varepsilon > 0$ and take the corresponding δ given by the equintegrability condition on r_n. Clearly w.l.o.g. $\delta \leq \varepsilon$. Consider $\phi \in C_b(Y)$ positive such that

$$\int \phi \xi \leq \delta^2/2.$$

From the weak convergence for n great enough

$$\int \phi \xi_n \leq \delta^2,$$

so that we can estimate

$$\int \phi r_n \xi_n \leq \int_{\phi > \delta} r_n \xi_n + \delta < \varepsilon + \delta.$$

Hence $\int \phi \zeta < 2\varepsilon$. □

Theorem 1.46. *Assume that the family of functions $\{r_n\} \subset L^1(m_n \otimes \mathcal{L}^1)$ given by (1.7.1) is such that*

$$(\mathbb{I}, \mathbb{I}, g_n)_\sharp \big(r_n m_n \otimes \mathcal{L}^1 \big) \rightharpoonup (\mathbb{I}, \mathbb{I}, g)_\sharp \zeta$$

with $\zeta \in \mathcal{P}(S \times \mathbb{R})$ and g being the ray map (Definition 1.25). Assume moreover

$$\forall T \geq 0 \ \forall \varepsilon > 0 \ \exists \delta > 0 \ \bigg(A \in \mathcal{B}(S \times [-T, T]), m_n \otimes \mathcal{L}^1(A) < \delta \quad \Longrightarrow \quad \int_A r_n m_n \otimes \mathcal{L}^1 < \varepsilon \bigg).$$

Then $\zeta = r m \otimes \mathcal{L}^1$ for some function $r \in L^1(m \otimes \mathcal{L}^1)$, measure $m \in \mathcal{P}(S)$ and the disintegration of μ is a.c. w.r.t. $\mathcal{H}^1$ on each geodesic.

Proof. Define for $k \in \mathbb{N}$

$$\phi_k \in C_c(\mathbb{R}),\ \phi_k \geq 0, \quad \phi_k(t) := \begin{cases} 1 & |t| \leq k, \\ 0 & |t| \geq k+1. \end{cases}$$

Let $\xi_{n,k} = m_n \otimes \mathcal{L}^1 \llcorner_{[-k-1,k+1]}$ and consider the functions $\tilde{r}_{n,k} := r_n(y,t)\phi_k(t)$. Since $m_n = (P_1)_\sharp(r_n m_n \otimes \mathcal{L}^1)$ and hence $m_n \rightharpoonup m = (P_1)_\sharp \zeta$, then

$$\xi_{n,k} \rightharpoonup m \otimes \mathcal{L}^1 \llcorner_{[-k-1,k+1]}$$

and the hypothesis of Proposition 1.45 are verified up to rescaling. So $\zeta = rm \otimes \mathcal{L}^1$.

The fact that $g_\sharp \zeta$ is a disintegration is a consequence of the a.c. of ζ along each geodesic: in this case the initial points have ζ-measure 0 and therefore g is invertible on a set of full μ-measure. □

In general the convergence of the graph of g_n is too strong: the next result considers a more general case.

Proposition 1.47. *Assume that $\tilde{\zeta} \in \Pi(rm \otimes \mathcal{L}^1, \mu)$ is concentrated on the graph of a Borel function $h : \mathcal{T} \times \mathbb{R} \to \mathcal{T}_e$ such that*

1. *$(y,t) \mapsto e(y) := f(h(y,t)) \in S$ is constant w.r.t. t,*
2. *it holds*
$$h(y,\cdot)_\sharp\big(r(y,\cdot)\mathcal{L}^1\big) \ll \mathcal{H}^1 \llcorner_{g(e(y),\mathbb{R})}.$$

Then the disintegration w.r.t. g has absolutely continuous conditional probability.

Proof. We can disintegrate the measure m as follows:

$$m = \int_S m_z (e_\sharp m)(dz),$$

and by the second assumption

$$h(y,\cdot)_\sharp(r(y,\cdot)\mathcal{L}^1) = g(e(y),\cdot)_\sharp(\tilde{r}(y,\cdot)\mathcal{L}^1),$$

for m-a.e. $y \in \mathcal{T}$. Hence by explicit computation,

$$\begin{aligned} \mu &= \int_S h(y,\cdot)_\sharp(r(y,\cdot)\mathcal{L}^1) m(dy) = \int_S g(e(y),\cdot)_\sharp(\tilde{r}(y,\cdot)\mathcal{L}^1) m(dy) \\ &= \int_S \bigg(\int_{e^{-1}(z)} g(z,\cdot)_\sharp(\tilde{r}(y,\cdot)\mathcal{L}^1) m_z(dy) \bigg) e_\sharp m(dz). \end{aligned}$$

To conclude the proof observe that

$$\begin{aligned} \int_{e^{-1}(z)} g(z,\cdot)_\sharp(\tilde{r}(y,\cdot)\mathcal{L}^1) m_z(dy) &= g(z,\cdot)_\sharp \bigg(\int_{e^{-1}(z)} \tilde{r}(y,\cdot)\mathcal{L}^1 m_z(dy) \bigg) \\ &= g(z,\cdot)_\sharp \bigg(\int_{e^{-1}(z)} \tilde{r}(y,\cdot) m_z(dy) \bigg) \mathcal{L}^1. \end{aligned}$$

□

Remark 1.48. Observe that some properties of r_n are preserved passing to the limit r. In relation with the previous section, we consider the following cases: for $A \subset X \times \mathbb{R}$ open

1. for some $\varepsilon > 0$
$$\big((r_n - \varepsilon) m_n \otimes \mathcal{L}^1\big)\llcorner_A \geq 0;$$

2. there exists $L > 0$ such that
$$r_n(y, \cdot) \in \mathrm{Lip}_L(A_y);$$

3. there exists $M > 0$ such that
$$TV(r_n(y, \cdot)\llcorner_A) \leq M.$$

The first condition yields that the assumptions of Corollary 1.44 holds in A. The second and third conditions imply that we are under the conditions for Remark 1.41 in A.

1.7.2 Approximations by metric spaces

In this section we explain a procedure to verify if the transport problem under consideration satisfies Assumption 2. The basic references for this sections are [23] and [26, 27].

We consider the following setting:

1. (X, d, d_L), $(X_n, d_n, d_{L,n})$, $n \in \mathbb{N}$, are metric structures satisfying the assumptions of page 19 and Remark 1.6: more precisely, $d_L, d_{L,n}$ l.s.c., $d_L \geq d, d_{L,n} \geq d_n$ and
$$\bigcup_{x \in K_1, y \in K_2} \gamma_{[x,y]}$$
is $d_n(d)$-compact if K_1, K_2 are $d_n(d)$-compact, $d_{L,n}(d_L)\llcorner_{K_1 \times K_2}$ uniformly bounded.

2. $\mu_n, \nu_n \in \mathcal{P}(X_n)$, $\mu_n \perp \nu_n$;

3. $\pi_n \in \Pi(\mu_n, \nu_n)$ is a $d_{L,n}$-cyclically monotone transference plan with finite cost.

For $\mu, \nu \in \mathcal{P}(X)$ let $\pi \in \Pi(\mu, \nu)$ be a generic transference plan.

Definition 1.49. We say that the structures $(X_n, d_n, d_{L,n}, \pi_n)$ *converge* to (X, d, d_L, π) if the following holds: there exists $C > 0$ such that for all $n \in \mathbb{N}$
$$\int d_{L,n} \pi_n \leq C$$
and there exist Borel sets $A_n \subset X_n$ and Borel maps $\ell_n : A_n \to X$ such that
$$(\ell_n \otimes \ell_n)_\sharp \pi_n \llcorner_{A_n \times A_n} \rightharpoonup \pi, \tag{1.7.2}$$
$$\big| d_L(\ell_n(x), \ell_n(y)) - d_{L,n}(x, y) \big| \leq 2^{-n}, \tag{1.7.3}$$
and if $(\ell_n(x_n), \ell_n(y_n)) \to (x, y)$, then
$$d_L(x, y) = \lim_n d_{L,n}(x_n, y_n). \tag{1.7.4}$$

As a first result, we show that also π is d_L-cyclically monotone with finite cost.

Proposition 1.50. *If $(X_n, d_n, d_{L,n}, \pi_n)$ converges to (X, d, d_L, π) and the plans π_n have uniformly bounded cost then also π has finite cost and is d_L-cyclically monotone.*

Proof. Since d_L is l.s.c.

$$\begin{aligned}\int d_L \pi \leq & \liminf_{n \to +\infty} \int d_L(\ell_n \otimes \ell_n)_\sharp \pi_n = \liminf_{n\to+\infty} \int d_L(\ell_n(x), \ell_n(y)) \pi_n(dxdy) \\ & \stackrel{(1.7.3)}{\leq} \liminf_{n\to+\infty} \bigg\{ \int d_{L,n}(x,y)\pi_n(dxdy) + 2^{-n} \bigg\} \leq C,\end{aligned}$$

for some $C < +\infty$.

Now let Γ_n be a $d_{L,n}$-cyclically monotone set with $\pi_n(\Gamma_n) = 1$: by standard regularity of Borel function and by Prokhorov Theorem we can assume that

1. Γ_n is σ-compact, $\Gamma_n = \cup_{m\in\mathbb{N}}\Gamma_{n,m}$ with $\Gamma_{n,m} \subset \Gamma_{n,m+1}$;
2. $(\ell_n \otimes \ell_n)(\Gamma_{n,m})$ is compact and $(\ell_n \otimes \ell_n)(\Gamma_{n,m}) \to \Gamma_m$ in the Hausdorff distance d_H;
3. $\pi_n(\Gamma_{n,m}) \geq 1 - 2^{-m}$.

It follows that: $\pi(\Gamma_m) \geq 1 - 2^{-m}$, hence

$$\pi\bigg(\bigcup_{m\in\mathbb{N}} \Gamma_m \bigg) = 1.$$

Since each Γ_m is the limit in Hausdorff distance of $(\ell_n \otimes \ell_n)(\Gamma_{n,m})$, (1.7.4) implies that Γ_m (and thus $\cup_m \Gamma_m$, because $\Gamma_m \subset \Gamma_{m+1}$) is d_L-cyclically monotone. $\square$

Note that since π is d_L-cyclically monotone, we can define the sets $\Gamma, \Gamma', G, G^{-1}, R, a, b$ of Section 1.2 as well as the quotient map f and the ray map g constructed in Section 1.3. The same sets and maps can be given for the structures $(X_n, d_n, d_{L,n})$: we will denote them with the subscript n.

For the transport problems in (X_n, d_n) with measures μ_n, ν_n, we assume the following.

Assumption 3 (Non degeneracy)**.** The $d_{L,n}$-cyclically monotone plan π_n satisfies Assumption 3 for all $n \in \mathbb{N}$.

This allows to write the disintegration of μ_n w.r.t. the ray equivalence relation R_n:

$$\mu_n = (g_n)_\sharp(r_n m_n \otimes \mathcal{L}^1) = \int g_n(y,\cdot)_\sharp(r_n(y,\cdot)\mathcal{L}^1) m_n(dy),$$

with $f_{n\,\sharp}\mu_n = m_n$ and $r_n \in L^1(m_n \otimes \mathcal{L}^1)$.

Lemma 1.51. *If $(X_n, d_n, d_{L,n}, \pi_n)$ converges to (X, d, d_L, π) then the structures $(S_n \times \mathbb{R}, \tilde{d}_n, \tilde{d}_{L,n}, \tilde{\pi}_n)$, where*

$$\tilde{d}_n = d_n \circ (g_n \otimes g_n), \quad \tilde{\pi}_n = (g_n^{-1} \otimes g_n^{-1})_\sharp(\pi_n), \quad \tilde{d}_{L,n}((y,t),(y',t')) = \begin{cases} |t - t'| & y = y' \\ +\infty & y \neq y', \end{cases}$$

converges to (X, d, d_L, π).

Proof. It is enough to observe that $\pi_n(G_n) = 1$, $\tilde{d}_{L,n} = d_{L,n} \circ (g_n \otimes g_n)$ on G_n and to replace the map ℓ_n with the map $\ell_n \circ g_n$. $\square$

By Lemma 1.51, in the following we assume that the ray map g_n is the identity map.

The next assumption is the fundamental one.

Assumption 4 (Equintegrability)**.** The L^1-functions r_n are equintegrable w.r.t. the measure $m_n \otimes \mathcal{L}^1$:

$$\forall \varepsilon > 0 \; \exists \delta > 0 \bigg((m_n \otimes \mathcal{L}^1)(A) < \delta \;\Rightarrow\; \int_A r_n m_n \otimes \mathcal{L}^1 < \varepsilon \bigg).$$

From now on we will assume that $(X_n, d_n, d_{L,n}, \pi_n) \to (X, d, d_L, \pi)$ in the sense of Definition 1.49, $(X_n, d_n, d_{L,n}, \pi_n)$ verifies Assumption 3 and Assumption 4 and $\int d_{L,n}\pi_n \leq C$.

Our aim is to prove that the structure (X, d, d_L, π) satisfies Assumption 2, which is equivalent to the fact that the marginal probabilities of the disintegration of μ w.r.t the ray equivalence relation R are a.c. w.r.t. $\mathcal{H}^1$.

The next lemma shows that in order to obtain our purpose we can perform some reductions without losing generality. We will write $\mu_k \nearrow \mu$ for $\mu_k \leq \mu_{k+1}$ and $\mu = \sup_k \mu_k$.

Lemma 1.52. *Let $\{\mu_k\}_{k\in\mathbb{N}} \subset \mathcal{M}(X)$, $\mu_k \geq 0$, be such that $\mu_k \nearrow \mu$ and assume that*

$$\mu_k = g_\sharp(r_k m_k \otimes \mathcal{L}^1), \quad r_k \geq 0,$$

where g is the ray map on $\mathcal{T}$. Then there exist $m \in \mathcal{P}(X)$, $r \in L^1(m \otimes \mathcal{L}^1)$, $r \geq 0$ such that the same formula holds for μ:

$$\mu = g_\sharp(rm \otimes \mathcal{L}^1).$$

Proof. Since $\int r_k(y,t)dt = 1$ it follows that $P_{1\,\sharp}(r_k m_k \otimes \mathcal{L}^1) = m_k$ and therefore $m_k \nearrow m$ with $m = f_\sharp\mu$ (recall that f is a section for the ray equivalence relation R). The convergence $\mu_k \nearrow \mu$ yields

$$\left(r_k \frac{dm_k}{dm}\right) m \otimes \mathcal{L}^1 \nearrow \zeta,$$

where $\mu = g_\sharp\zeta$. We conclude $\zeta = rm \otimes \mathcal{L}^1$ with $r := \sup_k r_k \frac{dm_k}{dm}$. □

A first reduction is given by the following lemma.

Lemma 1.53. *To prove that there exist $m \in \mathcal{P}(X)$, $r \in L^1(m \otimes \mathcal{L}^1)$, $r \geq 0$ such that*

$$\mu = g_\sharp(rm \otimes \mathcal{L}^1),$$

we can assume w.l.o.g. that there exist $\bar{x}, \bar{y} \in X$ and $q \geq 0$ such that

$$\pi\Big(\Big\{(x,y) : d(\bar{x},\bar{y}) > 8q, d(x,\bar{x}), d(y,\bar{y}) \leq q\Big\}\Big) = 1.$$

Moreover the d_L-cyclically monotone set Γ and the set of oriented transport rays G can be assumed to be compact subsets of $X \times X$.

Proof. Step 1. Since $\pi(\{x = y\}) = 0$ we can assume that $\Gamma \cap \{x = y\} = \emptyset$. Take two dense sequences $\{x_i\}_{i\in\mathbb{N}} \subset X$, $\{q_i\}_{i\in\mathbb{N}} \subset \mathbb{R}^+$ and consider the family of closed sets

$$\Gamma_{ijk} := \Big\{(x,y) : d(x_i, x_j) \geq 8q_k, d(x, x_i), d(y, x_j) \leq q_k\Big\}.$$

Then Γ_{ijk} is a countable covering of $X \times X \setminus \{x = y\}$.

Suppose now to have proven that for all $\mu_{ijk} = P_{1\,\sharp}(\pi\llcorner_{\Gamma_{ijk}})$ the disintegration formula holds with $\mathcal{H}^1$-a.c. marginal probabilities, then the same $\mathcal{H}^1$-a.c. property is true if we replace Γ_{ijk} with the finite union of sets $\Gamma_{i'j'k'}$.

Define

$$\tilde{\Gamma}_m := \bigcup_{n<m} \Gamma_{i_n j_m k_m},$$

where

$$\mathbb{N} \ni m \mapsto (i_m, j_m, k_m) \in \mathbb{N}^3$$

is a bijective map, and consider $\mu_m = P_{1\,\sharp}(\pi\llcorner_{\tilde{\Gamma}_m})$, then $\{\mu_m\}_{m\in\mathbb{N}}$ verifies the hypothesis of Lemma 1.52.

Step 2. It remains to show how to construct the approximating structure $\tilde{\pi}_n \in \mathcal{P}(X_n)$ converging in the sense of Definition 1.49 to $\pi \llcorner_{\Gamma_{ijk}}$. Since Γ_{ijk} is closed, there exists a sequence $\phi_l \in C_c(X \times X, [0,1])$ such that $\phi_l \searrow \chi_{\Gamma_{ijk}}$. Now $\phi_l \pi \searrow \pi \llcorner_{\Gamma_{ijk}}$ as $l \to +\infty$ and

$$\phi_l (\ell_n \otimes \ell_n)_\sharp \pi_n \rightharpoonup \phi_l \pi.$$

Hence there exists a subsequence $\{\phi_{l_i}(\ell_{n_i} \otimes \ell_{n_i})_\sharp \pi_{n_i}\}_{i \in \mathbb{N}}$ satisfying (1.7.2) with weak limit $\pi \llcorner_{\Gamma_{ijk}}$. If one defines

$$\tilde{\pi}_i = \big(\phi_{l_i} \circ \ell_{n_i}\big) \pi_{n_i},$$

then it is straightforward to show that $(X_i, d_i, d_{L,i}, \tilde{\pi}_i)$ converges to $(X, d, d_L, \pi \llcorner_{\Gamma_{ijk}})$ in the sense of Definition 1.49.

Step 3. Since Remark 1.19 yields that Γ, G are σ-compact, let $\Gamma = \cup_k \Gamma_k$, $G = \cup_k G_k$ with Γ_k, G_k compact and consider $\pi \llcorner_{\Gamma_k}$. The same reasoning done in Step 1 and Step 2. yields that it is enough to prove the a.c. of disintegration for $\pi \llcorner_{\Gamma_k}$. □

Therefore from now on we will assume that π is concentrated on the set

$$\Big\{(x,y) : d(\bar{x}, \bar{y}) > 8q, d(x, \bar{x}), d(y, \bar{y}) \leq q\Big\}.$$

Using the same reasoning of Lemma 1.53 one can also prove the following.

Lemma 1.54. *We can assume w.l.o.g. that the sets $A_n \subset X_n$ are compact and the maps $\ell_n : A_n \to X$ are continuous. Moreover $\ell_n(A_n)$ converges in Hausdorff distance to a compact set K on which μ and ν are concentrated.*

Proof. By Lusin Theorem and inner regularity of measures it follows that there exist $B_n \subset A_n$ such that

- $A_n \setminus B_n$ is compact;
- $\mu_n(B_n) \leq 1/n$;
- the map $\ell_n : A_n \setminus B_n \to X$ is continuous.

To prove the first part of the claim just observe that $(\ell_n \otimes \ell_n)_\sharp \pi_n \llcorner_{A_n \setminus B_n \times A_n \setminus B_n} \rightharpoonup \pi$.

The second part of the statement can be proven following the line of the second part of the proof of the Proposition 1.50. □

By Lemma 1.54 it is straightforward that for all n great enough we have

$$\big(\ell_{n\,\sharp} \mu_n\big)(B_{2q}(\bar{x})) = \big(\ell_{n\,\sharp} \nu_n\big)(B_{2q}(\bar{y})) = 1.$$

Lemma 1.55. *We can assume that the measure m_n is concentrated on a compact subset of*

$$\Big\{y \in S_n : \exists t, s > \delta : (y, -s), (y, 0), (y, t) \in P_{12}(\mathrm{graph}(\ell_n))\Big\}$$

for some fixed $\delta > 0$ and

$$\mu_n(S_n \times (-\infty, 0]) = 1, \quad \nu_n(S_n \times (4q, +\infty)) = 1. \tag{1.7.5}$$

Proof. Step 1. Defining

$$A_\delta := \Big\{(y,t) \in S_n \times \mathbb{R} : |a(y) - t| < \delta\Big\},$$

by Fubini Theorem

$$m_n \otimes \mathcal{L}^1(A_\delta) \leq \delta,$$

hence by Assumption 4, for any $\varepsilon > 0$ we can choose $\delta > 0$ such that $r_n m_n \otimes \mathcal{L}^1(A_\delta) < \varepsilon$ for al $n \in \mathbb{N}$.

Therefore we can assume that $r_n m_n \otimes \mathcal{L}^1$ is concentrated on a compact subset B_n of $\ell_n^{-1}(\bar{B}(\bar{x}, \frac{3}{2}q)) \setminus A_\delta$.

Step 2. Define the u.s.c. selection of B_n $t_n : S_n \to \mathbb{R}$ in the following way:

$$y \mapsto t_n(y) := \max \big\{ t \in \mathbb{R}, (y,t) \in B_n \big\}.$$

By removing a set of arbitrarily small measure we can assume that for all $y \in P_1(\mathrm{graph}(t_n))$ there exists $t > 4q$ such that

$$(y, t_n(y) + t) \in \ell_n^{-1}\bigg(\bar{B}\Big(\bar{y}, \frac{3}{2}q \Big) \bigg).$$

Step 3. The Borel transformation

$$B_n \ni (y,t) \mapsto (y, t - t_n(y))$$

maps $m_n \otimes \mathcal{L}^1$ into itself and in the new coordinates the section S_n satisfies the first part of the claim.

By the definition of G_n and $\mu_n \perp \nu_n$ it follows that μ_n and ν_n satisfy (1.7.5), see Remark 1.38. $\square$

Define the map

$$\begin{array}{rcl} h_n : S_n \times \mathbb{R} & \to & X \times \mathbb{R} \\ (y,t) & \mapsto & (\ell_n(y,0), t). \end{array}$$

and the measure $h_{n\,\sharp}(r_n m_n \otimes \mathcal{L}^1) = \tilde{r}_n \tilde{m}_n \otimes \mathcal{L}^1$, with $\tilde{m}_n = \ell_n(\cdot, 0)_\sharp m_n$.

Lemma 1.56. *The family of measures* $\{\tilde{r}_n \tilde{m}_n \otimes \mathcal{L}^1\}_{n\in\mathbb{N}} \subset \mathcal{P}(\ell_n(S_n \times \{0\}) \times \mathbb{R})$ *is tight and* $\tilde{r}_n$ *is equintegrable w.r.t.* $\tilde{m}_n \otimes \mathcal{L}^1$.

Proof. Performing the same calculation of (1.5.3)

$$C \geq \int d_{L,n} \pi_n = \int s \nu_n - \int s \tilde{r}_n \tilde{m}_n \otimes \mathcal{L}^1.$$

From (1.7.5), Lemma 1.55, it follows that $s \leq 0$, $\tilde{r}_n \tilde{m}_n \otimes \mathcal{L}^1$-a.e.. Hence $\tilde{r}_n \tilde{m}_n \otimes \mathcal{L}^1 \in \mathcal{P}(\ell_n(S_n \times \{0\}) \times (-\infty, 0]))$ and

$$0 \leq -\int s \tilde{r}_n \tilde{m}_n \otimes \mathcal{L}^1 \leq C,$$

therefore $\tilde{r}_n \tilde{m}_n \otimes \mathcal{L}^1$ is tight. Recall in fact that $\{S_n\}_{n\in\mathbb{N}}$ is a precompact sequence w.r.t. the Hausdorff distance by Lemma 1.54.

The equintegrability is straightforward:

$$\int_A \tilde{r}_n \tilde{m}_n \otimes \mathcal{L}^1 = \int_{(h_n)^{-1}(A)} r_n m_n \otimes \mathcal{L}^1$$

and $m_n \otimes \mathcal{L}^1((h_n)^{-1}(A)) = \tilde{m}_n \otimes \mathcal{L}^1(A)$. $\square$

Consider the following measure

$$\zeta_n := (h_n, \ell_n)_\sharp (r_n m_n \otimes \mathcal{L}^1) \in \Pi\Big(\tilde{r}_n \tilde{m}_n \otimes \mathcal{L}^1, (\ell_n)_\sharp(\mu_n) \Big) \in \mathcal{P}(X \times \mathbb{R} \times X).$$

Proposition 1.57. *Up to subsequences,* $\zeta_n \rightharpoonup \zeta$, *where* $\zeta \in \Pi(rm \otimes \mathcal{L}^1, \mu)$ *is supported on a Borel graph* $h : \mathcal{T} \times \mathbb{R} \to \mathcal{T}_e$ *such that* $t \mapsto h(y,t)$ *is the* d_L *1-Lipschitz curve* $R(y)$ *for* m*-a.e.* $y \in X$.

Proof. Step 1. The convergence to the correct marginals is a consequence of (1.7.2)

$$(P_2)_\sharp \zeta_n = (\ell_n \circ g_n)_\sharp (r_n m_n \otimes \mathcal{L}^1) = (\ell_n)_\sharp \mu_n \rightharpoonup \mu,$$

and by Lemma 1.56

$$(P_1)_\sharp \zeta_n = \tilde{r}_n \tilde{m}_n \otimes \mathcal{L}^1 \rightharpoonup rm \otimes \mathcal{L}^1.$$

Step 2. Since up to subsequence $\zeta_n \rightharpoonup \zeta$, using the same technique of Lemma 1.50, we can assume that $K_n := (h_n, \ell_n)(S_n \times \mathbb{R})$ is compact and $d_H(K_n, \mathrm{graph}(h)) \to 0$ where $\mathrm{graph}(h)$ is a compact set supporting ζ and h is the associated multivalued function.

Step 3. Let $(y,t,x) \in \mathrm{graph}(h)$, then by the definition of convergence in the Hausdorff metric, there exists a sequence $(\ell_n(y_n, 0), t_n, \ell_n(y_n, t_n)) \to (y,t,x)$. Hence from

$$d_{L,n}\big((y_n, t_n), (y_n, 0)\big) = |t_n| \to |t|,$$

we deduce by (1.7.4) that $d_L(x,y) = |t|$. In particular this implies that if $t = 0$ then $x = y$.

Step 4. Let $(y,t,x), (y,t',x') \in \mathrm{graph}(h)$ with $t < 0$ and $t' > 0$. Again by the Hausdorff convergence there exist two sequences satisfying

$$\big(\ell_n(y_n, 0), t_n, \ell_n(y_n, t_n)\big) \to (y,t,x), \quad \big(\ell_n(y'_n, 0), t'_n, \ell_n(y'_n, t'_n)\big) \to (y,t',x').$$

Since $d_L(y,y) = 0$, from (1.7.4) we deduce

$$d_{L,n}\big((y_n, 0), (y'_n, 0)\big) \to 0$$

hence by the definition of $d_{L,n}$, for n great enough $y_n = y'_n$. Therefore

$$d_L(x,x') = \lim_{n \to +\infty} d_{L,n}((y_n,t_n),(y'_n,t'_n)) = |t| + t',$$

and by Step 3 we conclude that $d_L(x,x') = d_L(x,y) + d_L(y,x')$.

Step 5. Let $(y,t,x) \in \mathrm{graph}(h)$: we now show that

$$t \geq 0 \;\Rightarrow\; (y,x) \in G, \quad -t \geq 0 \;\Rightarrow\; (y,x) \in G^{-1}.$$

We will prove only the first implication for $t > 0$. Since following Lemma 1.54 we can take G_n compact such that

1. $(\ell_n \otimes \ell_n)(G_n) \to \hat{G}$ in the Hausdorff metric;
2. $\hat{G} \subset G$,

it is enough to show that there exists a sequence $(\ell_n(y_n,0), t_n, \ell(y_n,t_n)) \to (y,t,x)$ so that $(y_n, x_n) \in G_n$ for all n, but this last implication is straightforward.

Step 6. We next show that for any $y \in P_1(\mathrm{graph}(h))$ there exist $t_-, t_+ \geq \delta$ and x_-, x_+ such that $(y, -t_-, x_-), (y, t_+, x_+) \in \mathrm{graph}(h)$. In fact we recall that for all $y_n \in S_n$ there exist $t_{-,n}, t_{+,n} \geq \delta$, for some strictly positive constant δ, such that

$$\big((y_n, -t_{-,n}), (y_n, 0)\big), \big((y_n, 0), (y_n, t_{+,n})\big) \in G_n.$$

Hence chose $y_n \in S_n$ such that $\ell_n(y_n) \to y$ and pass to converging subsequences to obtain the claim.

Step 7. Since for $y \in P_1(\mathrm{graph}(h))$ there exist x, x' such that $(x,y), (y,x') \in G \setminus \{x = y\}$, then $(x,x') \in G$, $y \in \mathcal{T}$ and h is single valued. The same computation of Point 5 yields that

$$\big\{(y, h(t,y)), t \geq 0\big\} \cup \big\{(h(t,y), y), t \leq 0\big\} \subset G,$$

and from this it follows that $h(y, \mathbb{R}) \subset R(y)$.

Again from Point 5 one obtains that $d_L(y, h(t,y)) = |t|$ and therefore $t \mapsto g^{-1}(h(y,t)) = g^{-1}(y) + t$. □

Theorem 1.58. *Let $(X_n, d_n, d_{L,n}, \pi_n) \to (X, d, d_L, \pi)$ and $(X_n, d_n, d_{L,n}, \pi_n)$, $n \in \mathbb{N}$, verifies Assumption 3 and Assumption 4. Then the marginal measure $\mu = P_{1\,\sharp}(\pi)$ satisfies Assumption 2.*

Proof. The measure ζ constructed in the Proposition 1.57 satisfies the hypothesis of Proposition 1.47. Therefore the marginal probabilities of the disintegration of μ are absolutely continuous with respect to $\mathcal{H}^1$ and therefore μ verifies Assumption 2. □

Remark 1.59. As in Remark 1.48, if we know more regularity of the disintegrations for the approximating problems, we can pass them to the limit. Here the key observation is that geodesics converge to geodesics, so uniform continuous functions on them converge pointwise to continuous functions.

A special case is when $d_L = d$: a natural approximation is by transport plans where ν is atomic, with a finite number of atoms. This case can be studied with more standard techniques, we refer to the analysis contained in [10].

1.8 Monge problem for curved metric measure space

In this section we recall the definition of Measure Contraction Property (MCP) and then we prove that for a metric measure space (X, d, η) satisfying MCP, the Monge minimization problem with marginal measures μ and ν with $\mu \ll \eta$ and cost d admits a solution. We show moreover that the hypotheses of Corollary 1.44 hold, and if suppμ and suppν are at positive distance then the assumptions of Lemma 1.40 are satisfied, i.e. the current $\dot{g}$ is normal.

The main reference for this section is [24].

From now on $d = d_L$ and $\eta \in \mathcal{M}^+(X)$ is a locally finite measure on X. Since $d_L = d$ there exists a Lipschitz function φ potential for the transport problem: hence in the following we will set

$$\Gamma = \Gamma' = G = \Big\{(x, y) \in X \times X : \varphi(x) - \varphi(y) = d(x, y)\Big\},$$

where ϕ is a potential for the transport problem.

Let H be the set of all geodesics: we regard H as a subset of $\mathrm{Lip}_1([0, 1], X)$ with the uniform topology. Define the evaluation map $e_t(\gamma)$ by

$$\begin{array}{rcl} e : [0,1] \times H & \to & X \\ e_t(\gamma) & \mapsto & \gamma(t) \end{array} \tag{1.8.1}$$

It is immediate to see that $e_t(\gamma)$ is continuous.

A *dynamical transference plan* Ξ is a Borel probability measure on H, and the path $\{\xi_t\}_{t\in[0,1]} \subset \mathcal{P}^2(X)$ given by $\xi_t = (e_t)_\sharp \Xi$ is called *displacement interpolation* associated to Ξ. We recall that $\mathcal{P}^2(X)$ is the set of Borel probability measures ξ satisfying $\int_X d^2(x, y)\xi(dy) < \infty$ for some (and hence all) $x \in X$.

Define for $K \in \mathbb{R}$ the function $s_K : [0, +\infty) \to \mathbb{R}$ (on $[0, \pi/\sqrt{K})$ if $K > 0$)

$$s_K(t) := \begin{cases} (1/\sqrt{K}) \sin(\sqrt{K}t) & \text{if } K > 0, \\ t & \text{if K} = 0, \\ (1/\sqrt{-K}) \sinh(\sqrt{-K}t) & \text{if } K < 0, \end{cases} \tag{1.8.2}$$

and let $N \in \mathbb{N}$.

Definition 1.60. A metric measure space (X, d, η) is said to satisfies the (K, N)-*measure contraction property* ($MCP(K, N)$) if for every point $x \in X$ and η-measurable set $A \subset X$ with $\eta(A) > 0$ there exists a displacement interpolation $\{\xi_t\}_{t\in[0,1]}$ associated to a dynamical transference plan $\Xi = \Xi_{x,A}$ satisfying the following:

1. We have $\xi_0 = \delta_x$ and $\xi_1 = \eta(A)^{-1}\eta_{\llcorner A}$;
2. for $t \in [0, 1]$

$$\eta \geq (e_t)_\sharp \bigg(t \bigg\{ \frac{s_K(td(x, \gamma(1)))}{s_K(d(x, \gamma(1)))} \bigg\}^{N-1} \eta(A)\Xi \bigg),$$

where we set $0/0 = 1$.

From now on we will assume the metric measure space (X, d, η) to satisfies $MCP(K, N)$ for some $K \in \mathbb{R}$ and $N \in \mathbb{N}$. Recall that $MCP(K, N)$ implies that (X, d) is locally compact, Lemma 2.4 of [24].

The strategy to prove Assumption 2 for any d-cyclically monotone plan is the following: first we prove that for any $\pi \in \Pi(\mu, \delta_x)$ d-monotone with x arbitrary, the marginal probabilities of η obtained by the disintegration induced by the ray map g are absolutely continuous w.r.t. $\mathcal{H}^1$ and their densities satisfy some uniform estimates. Then we observe that these estimates hold true also for any $\pi \in \Pi(\mu, \sum_{i \leq I} c_i \delta_{x_i})$ d-monotone. Finally we show that the same estimates hold for general transference plans and therefore we deduce that the densities of the marginals obtained by disintegrating η w.r.t. any d-monotone plan π are absolutely continuous w.r.t. $\mathcal{H}^1$.

By Lemma 1.52, it is enough to assume that there exists $K_1, K_2 \subset X$ compact set, such that $\mu(K_1) = \nu(K_2) = 1$ and $d_H(K_1, K_2) < +\infty$. Hence we can assume that $\operatorname{diam}(X) < +\infty$ and $\eta(X) = 1$.

Lemma 1.61. *Consider $\bar{x} \in X$ and let $\pi \in \Pi(\mu, \delta_{\bar{x}})$ be the unique d-cyclically monotone transference plan. Then η and the optimal flow induced by π verify Assumption 2: more precisely, $\eta = g_\sharp(qm \otimes \mathcal{L}^1)$ and the density q satisfies the estimate*

$$q(y,t) \geq \left\{ \frac{s_K(d(g(y,t),\bar{x}))}{s_K(d(g(y,s),\bar{x}))} \right\}^{N-1} q(y,s) \tag{1.8.3}$$

for m-a.e. $y \in S$, for any $s \leq t$ such that $d(g(y,t),x) > 0$.

We recall that S is a section for the ray equivalence relation. Since $\mu \ll \eta$, (1.8.3) implies that $\mu = g_\sharp(rm \otimes \mathcal{L}^1)$ with $r \leq q$.

Proof. First observe that the potential for the transport problem is

$$\varphi(x) := \varphi(\bar{x}) + d(x, \bar{x}),$$

so that the geodesics used by π are exactly $H_{\bar{x}} := H \cap e_0^{-1}(\bar{x})$, in the sense that

$$G = \Big\{ \big(\gamma(1-s), \gamma(1-t)\big), s \leq t, \gamma \in H_{\bar{x}} \Big\}.$$

Step 1. We first prove that the set of initial points $A = a(X)$ has η-measure zero. Suppose by contradiction that $\eta(A) > 0$ and let $\Xi_{\bar{x},A}$ be the dynamical transference plan associated: we can assume that $\Xi_{\bar{x},A}$ is supported on the set $H_{\bar{x},A} := H_{\bar{x}} \cap e_1^{-1}(A)$. Then the evolution of A by the geodesics of $H_{\bar{x},A}$ can be defined as

$$A^s := e_{1-s}(H_{\bar{x},A}).$$

By Condition 2 of Definition 1.60 and the fact that $e_{1-s}^{-1}(A^s) = H_{\bar{x},A}$

$$\eta(A^s) \geq \eta(A) \int_{H_{\bar{x},A}} (1-s) \left\{ \frac{s_K((1-s)d(\bar{x},\gamma(1)))}{s_K(d(\bar{x},\gamma(1)))} \right\}^{N-1} \Xi_{x,A}(d\gamma) > 0, \tag{1.8.4}$$

for all $s \in [0,1)$. Since all A^s are disjoint being the space non branching, it follows that $\eta(A) = 0$.

Step 2. For A with $\eta(A) > 0$ let $\Xi_{\bar{x},A}$ be the dynamical transference plan concentrated on a set $H_{\bar{x},A} := H_{\bar{x}} \cap e_1^{-1}(A)$. Denote as before $A^s := e_{1-s}(H_{\bar{x},A})$.

Observe that since the set initial point has η-measure zero, we can disintegrate η w.r.t. the ray equivalence relation: using the disintegration formula $\eta = \int \eta_y m(dy)$ the same estimate as in (1.8.4) yields

$$\int \eta_y(A^s) m(dy) \geq \int \eta_y(A) m(dy) \bigg(\int_{H_{\bar{x},A}} (1-s) \left\{ \frac{s_K((1-s)d(\bar{x},\gamma(1)))}{s_K(d(\bar{x},\gamma(1)))} \right\}^{N-1} \Xi_{\bar{x},A}(d\gamma) \bigg).$$

By evaluating the above formula on sets of the form $A = g(S \times [t_1, t_2])$, where g is the ray map such that $g(y, 0) = \bar{x}$ for all y, gives

$$\begin{aligned}
&\int_S \eta_y\big(g(y, [t_1, t_2](1-s))\big) m(dy) \\
&\quad \geq \int_S \eta_y\big(g(y, [t_1, t_2])\big) m(dy) \bigg(\int_{H_{\bar{x},A}} (1-s) \bigg\{ \frac{s_K((1-s) d(\bar{x}, \gamma(1)))}{s_K(d(\bar{x}, \gamma(1)))} \bigg\}^{N-1} \Xi_{\bar{x},A}(d\gamma) \bigg) \\
&\quad \geq \int_S \eta_y\big(g(y, [t_1, t_2])\big) m(dy) \min_{c \in [t_1, t_2]} \bigg\{ (1-s) \frac{s_K((1-s)|c|)}{s_K(|c|)} \bigg\}^{N-1}.
\end{aligned}$$

and therefore for m-a.e. y and every t_1, t_2

$$\eta_y\big(g(y, [t_1, t_2](1-s))\big) \geq \eta_y\big(g(y, [t_1, t_2])\big) \min_{c \in [t_1, t_2]} \bigg\{ (1-s) \frac{s_K((1-s)|c|)}{s_K(|c|)} \bigg\}^{N-1}. \tag{1.8.5}$$

Step 3. For $t_1 < 0$ consider the family of disjoint open sets

$$t_1 \bigg(1 - \frac{k}{2n}, 1 - \frac{k+1}{2n} \bigg), \quad k = \{0, 1, \ldots, n-1\}.$$

The above estimate and the fact that η_y is probability yield

$$\eta_y \bigg\{ g\bigg(y, t_1 \bigg(1, 1 - \frac{1}{2n} \bigg) \bigg) \bigg\} \leq \frac{1}{n} \max_{c \in [t_1, t_1/2]} \bigg\{ 2 \frac{s_K(|c|)}{s_K(2|c|)} \bigg\}^{N-1}.$$

Hence $\eta_y = q \mathcal{H}^1 \llcorner_{g(y, \mathbb{R})}$ and q satisfies (1.8.3). $\square$

Lemma 1.62. *Let $\pi \in \Pi(\mu, \sum_{i \leq I} c_i \delta_{x_i})$ d-cyclically monotone. Then the conditional probabilities of the disintegration of η w.r.t. the ray equivalence relation induced by π are absolutely continuous w.r.t. $\mathcal{H}^1$ and the density $q(y, \cdot)$ satisfies*

$$q(y, t) \geq \bigg\{ \frac{s_K(d(g(y, t), b(y)))}{s_K(d(g(y, s), b(y)))} \bigg\}^{N-1} q(y, s).$$

Proof. Let φ be a potential for the transport problem with marginal μ and ν. Define

$$E_i := \Big\{ z \in \mathcal{T}_e : \varphi(z) - \varphi(x_i) = d(z, x_i) \Big\}.$$

Now each E_i is sent by the optimal geodesic flow to x_i, so we can perform exactly the same calculations done in Lemma 1.61. Indeed $E_i \cap E_j \subset a(X)$ which has η-measure zero, $\eta_{\llcorner E_i}$ verifies (2) of Definition 1.60 along the geodesic flow connecting E_i to x_i. $\square$

Given $\tilde{H} \subset \mathrm{Lip}_1([0, 1], X)$ a set of geodesics and $A \subset X$, define

$$A^{s, \tilde{H}} := e_{1-s}(e_1^{-1}(A) \cap \tilde{H}). \tag{1.8.6}$$

Lemma 1.63. *Assume that there exists two compact sets $K_1, K_2 \subset X$ such that*

1. *$\mu(K_1) = \nu(K_2) = 1$;*
2. *there exist $0 < a \leq b < +\infty$ such that*

$$a = \min_{x_1 \in K_1, x_2 \in K_2} d(x_1, x_2) \leq \max_{x_1 \in K_1, x_2 \in K_2} d(x_1, x_2);$$

3. K_2 is a section of R.

Then if

$$H(G) := \Big\{\gamma \in H : \exists y \in K_2 \Big(\varphi(\gamma(0)) - \varphi(\gamma(1)) = d(\gamma(0), \gamma(1)) \wedge \gamma(0) = y\Big)\Big\}, \tag{1.8.7}$$

where φ is the potential for the transport problem with marginal μ and ν, then

$$\eta(K_1^{s,H(G)}) \geq \eta(K_1) \min_{a\leq c\leq b} \bigg\{(1-s)\frac{s_K((1-s)c)}{s_K(c)}\bigg\}^{N-1}.$$

Proof. Step 1. It follows directly from Lemma 1.62 that the statement holds for $\nu = \sum_{i\leq I} c_i\delta_{y_i}$.

We thus consider the sequence of approximating problem constructed as follows: let $\{y_i\}_{i\in\mathbb{N}}$ be a dense sequence in K_2 and for $I \in \mathbb{N}$ define

$$\varphi_I(x) := \min\Big\{\varphi(y) + d(x,y), y \in \{y_1, \cdots, y_I\}\Big\},$$

$$E_{i,I} := \Big\{x \in X : \varphi_I(x) - \varphi_I(y_i) = d(x, y_i), i \leq I\Big\},$$

$$\nu_I = \sum_{i\leq I} c_{i,I}\delta_{y_i}, \quad \text{where} \quad c_{i,I} = \mu\bigg(E_{i,I} \setminus \bigcup_{j\neq i} E_{j,I}\bigg).$$

Clearly φ_I is a potential for the transport problem with marginal μ and ν_I and let

$$H(G_I) := \Big\{\gamma \in H : \varphi_I(\gamma(0)) - \varphi_I(\gamma(1)) = d(\gamma(0), \gamma(1)) \wedge \gamma(0) \in \{y_1, \cdots, y_I\}\Big\}.$$

Step 2. Observe that $K_1^{s,H(G)}$ is compact. In fact, since K_1 and K_2 are compact, $H(G) \cap e_1^{-1}(K_1)$ is compact and since e_{1-s} is continuous $K_1^{s,H(G)} = e_{1-s}(H(G))$ is compact. For the same reasons the sets $K_1^{s,H(G_I)}$ are compact.

Step 3. $K_1^{s,H(G_I)}$ is contained in a compact set and $\varphi_I \to \varphi$ as $I \to +\infty$, so that up to subsequences $K_1^{s,H(G_I)}$ converges in Hausdorff distance to a compact subset of $K_1^{s,H(G)}$. By the upper semicontinuity of Borel bounded measures with respect to Hausdorff convergence for compact sets the claim follows. □

Theorem 1.64. *If $\pi \in \Pi(\mu, \nu)$ d-monotone then $\eta\llcorner_{\mathcal{T}_e} = g_\sharp(qm \otimes \mathcal{L}^1)$, where $\mathcal{T}_e$ is the transport set with end points* (1.2.5b)*, and for m-a.e. y and $s \leq t$ it holds*

$$\bigg\{\frac{s_K(d(g(y,t), b(y)))}{s_K(d(g(y,s), b(y)))}\bigg\}^{N-1} \leq \frac{q(y,t)}{q(y,s)} \leq \bigg\{\frac{s_K(d(g(y,t), a(y)))}{s_K(d(g(y,s), a(y)))}\bigg\}^{N-1} \tag{1.8.8}$$

Proof. Step 1. We first show that the set of initial points has η-measure zero. In fact suppose by contradiction that $\eta(a(S)) > 0$, where S is a section for the ray equivalence relation of π. Hence we can assume that S and $a(S)$ are compact and at strictly positive distance.

Applying Lemma 1.63 to the transport problem with marginals $\eta\llcorner_{a(S)}$ and $f_\sharp\eta$, where f is the quotient map, it follows that $\eta(a(S)) = 0$.

Step 2. Since the initial points have η-measure zero, we can disintegrate $\eta\llcorner_{\mathcal{T}_e}$ w.r.t. the ray equivalence relation obtaining $\eta\llcorner_{\mathcal{T}_e} = \int \eta_y m(dy)$. By a standard covering argument, it is enough to prove the statement on the set

$$D_\varepsilon := \big\{x : d(x, b(x)) \geq \varepsilon\big\}.$$

For any $0 < \delta < \varepsilon$ we can take the section S compact such that $d(f(x), b(x)) = \delta$, in particular we have $g(y, \delta) = b(y)$.

For $S' \subset S$ and $t_1 < t_2$ consider $\eta\llcorner_{g^{-1}(S'\times[t_1,t_2])}$. Applying Lemma 1.63 with

$$\mu = \frac{\eta\llcorner_{g^{-1}(S'\times[t_1,t_2])}}{\eta(g^{-1}(S' \times [t_1, t_2]))}, \quad \nu = f_\sharp\mu$$

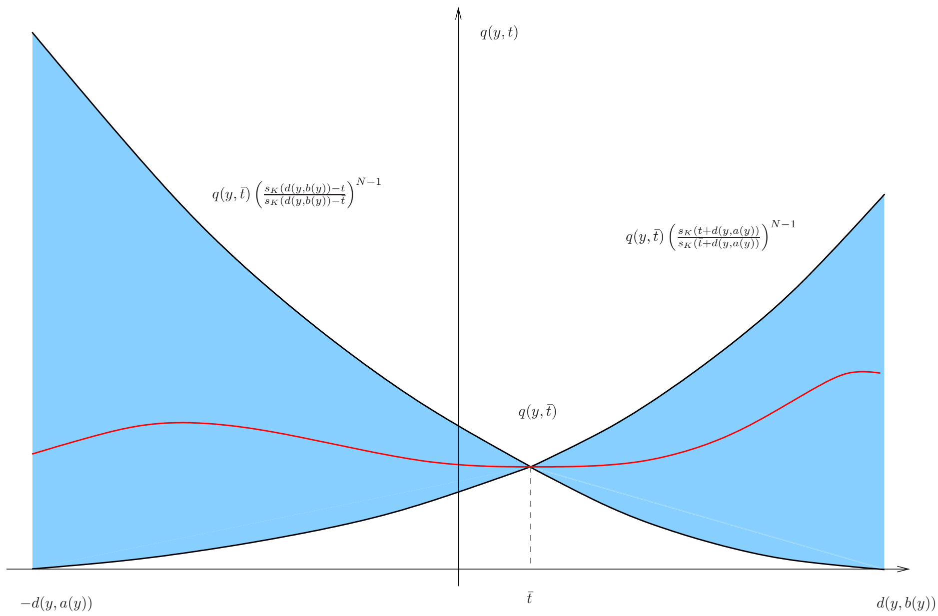

Figure 1.4: The region where $q(y,t)$ takes values.

where f is the quotient map for the ray equivalence relation R, it holds

$$\int_{S'} \eta_y\big(g(y,[t_1,t_2](1-s))\big)m(dy) \geq \min_{c\in[t_1,t_2]} \left\{(1-s)\frac{s_K\big((1-s)|c|\big)}{s_K\big(|c|\big)}\right\}^{N-1} \int_{S'} \eta_y\big(g(y,[t_1,t_2])\big)m(dy).$$

As in Step 2 of the proof of Lemma 1.61, the estimate (1.8.5) holds for m-a.e. y and every $t_1 < t_2$ and we deduce

$$\left\{\frac{s_K\big(d(g(y,t),b(y))-\delta\big)}{s_K\big(d(g(y,s),b(y))-\delta\big)}\right\}^{N-1} \leq \frac{q(y,t)}{q(y,s)}.$$

Letting $\delta \to 0$, we obtain the left hand side of (1.8.8).

Step 3. The right hand side of of (1.8.8) is obtained by the same procedure taking

$$F_\varepsilon := \big\{x : d(d,a(x)) \geq \delta\big\}$$

and the section S such that $d(y,a(y)) = \delta$ for all $y \in S$. □

Since $\mu \ll \eta$, it follows that also the densities of the conditional probabilities of μ are absolutely continuous w.r.t. $\mathcal{H}^1$, and therefore we have the following corollary.

Corollary 1.65. *Let (X,d,η) satisfies $MCP(K,N)$, let $\mu,\nu \in \mathcal{P}(X)$ with $\mu \ll \eta$, then there exists a μ-measurable map $T : X \to X$ such that $T_\sharp\mu = \nu$ and*

$$\int d(x,T(x))\mu(dx) = \min_{\pi\in\Pi(\mu,\nu)} \int d(x,y)\pi(dxdy).$$

We can obtain additional regularity of the conditional probabilities η_y under $MCP(K,N)$: in particular we deduce that the conclusion of Corollary 1.44 holds and if the support of μ and ν are compact sets with empty intersection the statements of Lemma 1.40 and Remark 1.41 are true.

Lemma 1.66. *The marginal densities*

$$\big(-d(a(y),y),d(y,b(y))\big) \ni t \mapsto q(y,t) \in \mathbb{R}^+$$

are strictly positive Lipschitz continuous for m-a.e. $y \in S$, and for some constant $C > 0$

$$\mathrm{Tot.Var.}\big(q(y,\cdot)\big) \leq \frac{C}{d(a(y),b(y))}.$$

Proof. From (1.8.8) it follows immediately that the function $q(y,t) > 0$ and Lipschitz continuous for $t \in (-d(y,a(y)), d(y,b(y)))$ and m-a.e. y. By differentiating it follows that

$$-(N-1)\frac{s_K'(d(g(y,t),b(y)))}{s_K(d(g(y,t),b(y)))} \leq \frac{q'(y,t)}{q(y,t)} \leq (N-1)\frac{s_K'(d(g(y,t),a(y)))}{s_K(d(g(y,t),a(y)))}. \tag{1.8.9}$$

In particular $q(y,\cdot)$ is Lipschitz.

For notational convenience let us assume that $d(a(y),y) = d(y,b(y)) = l$. From (1.8.8) one can prove that

$$q(y,t) \geq q(y,0) \cdot \begin{cases} \dfrac{s_K(l-t)}{s_K(l)}, & t \geq 0 \\ \dfrac{s_K(-l+t)}{s_K(-l)}, & t \leq 0 \end{cases}$$

Since $\int q(y,t)dt = 1$ it follows that

$$q(y,0) \leq c_K(d(a(y),b(y))),$$

where

$$c_K(t) := \frac{s_k(t/2)^{N-1}}{2}\bigg(\int_0^{t/2} s_K(\tau)^{N-1}d\tau\bigg)^{-1} \leq \frac{C}{t},$$

being C a constant depending only on K.

To show that

$$\int_{-l}^{l} |q'(y,t)|dt < +\infty,$$

it is enough to prove

$$\int_{-l}^{0} |q'(y,t)|dt < +\infty,$$

From (1.8.9) it follows

$$\omega'(y,t) := q'(y,t) + (N-1)\frac{s_K'(l-t)}{s_K(l)}q(y,0) \geq 0$$

so that

$$\mathrm{Tot.Var.}\big(\omega(y,\cdot)\big) \leq \bigg(1 + (N-1)\bigg(\frac{s_K(2l)}{s_K(l)} - 1\bigg)\bigg)q(y,0).$$

Hence

$$\begin{aligned}
\mathrm{Tot.Var.}\big(q(y,\cdot),(-\ell,0]\big) &\leq \mathrm{Tot.Var.}\big(\omega(y,\cdot),(-\ell,0]\big) + \mathrm{Tot.Var.}\bigg((N-1)\frac{s_K'}{s_K},(-\ell,0]\bigg)q(0,y) \\
&\leq \mathrm{Tot.Var.}\big(\omega(y,\cdot),(-\ell,0]\big) + (N-1)\frac{s_K'(2l)}{s_K(l)}q(0,y) \\
&\leq \bigg(1 + 2\bigg(\frac{s_K(2l)}{s_K(l)} - 1\bigg)q(y,0).
\end{aligned}$$

Collecting all the estimates, we get

$$\mathrm{Tot.Var.}\big(q(y,\cdot)\big) \leq 2\bigg(1 + 2\bigg(\frac{s_K(2l)}{s_K(l)} - 1\bigg)c_K(2l).$$

$\square$

In general, the current $\dot{g}$ is not normal, as one can easily verify in $\mathbb{T}^2$ with the standard distance.

1.9 Examples

We end this chapter with some examples which shows how the different hypotheses of Section 1.1 enter into the analysis. In the following we denote the standard Euclidean scalar product in $\mathbb{R}^d$ as $\cdot$ and the standard distance in $\mathbb{T}^d$ by $|\cdot|$. We will also denote points by $p = (x, y, z, \dots) \in \mathbb{R}^d$, and α a fixed constant in $[0, 1] \setminus \mathbb{Q}$.

Example 1 (Non strongly consistent disintegration along rays)**.** Consider the metric space

$$(X, d) = \big(\mathbb{T}^2, |\cdot|\big)$$

and the l.s.c. distance in the local chart $X = \{(x, y) : 0 \leq x, y < 1\}$

$$d_L(p_1, p_2) := \begin{cases} |x_1 - x_2 + i| & y_1 - y_2 = \alpha(x_1 - x_2) + i\alpha + n \\ +\infty & \text{otherwise} \end{cases}$$

for $i, n \in \mathbb{Z}$. The sets D_L are given by

$$D_L(p_1) = \Big\{(x, y) : y = y_1 + \alpha(x - x_1 + i) \mod 1, i \in \mathbb{N}\Big\},$$

so that it is easy to see that the partition $\{D_L(p)\}_{p \in X}$ does not yield a strongly consistent disintegration. Since $t \mapsto (t \mod 1, \alpha t \mod 1)$ is a continuous not locally compact geodesic, Condition (5) is not verified in this system.

Consider the measures $\mu = \mathcal{L}^2 \llcorner \mathbb{T}$ and the map $T : (x, y) \mapsto (x, y + \alpha \mod 1)$: being μ invariant w.r.t. translations, one has $T_\sharp \mu = \mu$, and moreover

$$\int d_L(x, T(x)) \mu(dx) = 1.$$

If we consider points $(p_i, (x_i, y_i + \alpha \mod 1))$, $i = 1, \dots, I$, then the only case for which $d_L(p_{i+1}, p_i) < +\infty$ is when $p_{i+1} = (x_i + t \mod 1, y_i + \alpha t \mod 1)$ for some $t \in \mathbb{R}$, i.e. they belong to the geodesic

$$\mathbb{R} \ni t \mapsto (x_i + t \mod 1, y_i + \alpha t \mod 1) \in X.$$

Hence, to prove d_L-cyclical monotonicity, it is sufficient to consider path which belongs to a single geodesic, where d_L reduces to the the one dimensional length:

$$d_L\big((x, y), (x + t \mod 1, y + \alpha t \mod 1)\big) = |t|.$$

Since translations in $\mathbb{R}$ are cyclically monotone w.r.t. the absolute value, we conclude that T is d_L-cyclically monotone.

The fact that the optimal rays coincide with the sets D_L yields that the disintegration is not strongly consistent, in particular there is not a Borel section up to a saturated negligible set. Note that every transference plan which leaves the common mass in the same place has cost 0, so that this example shows the necessity of Condition (5) for Proposition 1.27.

Example 2 (Non optimality of transport map)**.** Consider the space $\mathbb{R}^2$ with the distance

$$M(x, y) := \begin{bmatrix} m(y) & 0 \\ 0 & 1 \end{bmatrix}.$$

The basic assumption is that

1. m is a symmetric bell shaped function such that $ym'(y) > 0$ for all $y \neq 0$,
2. $m(0) = 1$, $m(y) = (1 + y^2)^{-1}$ for $|y| \leq 1$

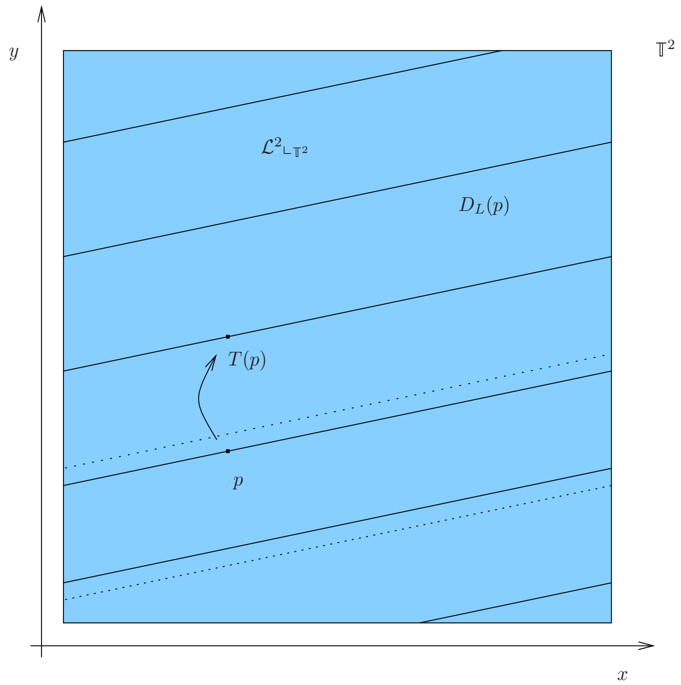

Figure 1.5: The metric space of Example 1

3. $m \geq (1+y^2)^{-1}$ and

$$\lim_{|y| \to +\infty} m(y) = \frac{1}{4}.$$

The only non zero Christoffel symbols of the Levi-Civita connection are

$$\Gamma^1_{12} = \Gamma^1_{21} = \frac{m'}{2m}, \quad \Gamma^2_{11} = -\frac{m'}{2},$$

and the equation for geodesics can be computed explicitly to be

$$\ddot{x} + \frac{m'(y)}{m(y)} \dot{x}\dot{y} = 0, \quad \ddot{y} - \frac{m'(y)}{2} (\dot{x})^2 = 0.$$

The first can be integrated into

$$\dot{x} m(y) = C,$$

and substituting into the second we obtain

$$\ddot{y} - \frac{C^2 m'(y)}{2m^2(y)} = 0, \quad \dot{y}^2 + \frac{C^2}{m(y)} = D.$$

In the case $m(y) = (1+y^2)^{-1}$ we obtain the explicit solution

$$\ddot{y} + C^2 y = 0, \quad y = a\sin(Ct) + b\sin(Ct).$$

$$x = x_0 + C\left(1 + \frac{a^2+b^2}{2}\right)t + \frac{a^2-b^2}{4}\sin(2Ct) - \frac{ab}{2}\cos(2Ct).$$

In particular, if the initial point is $(0,0)$ an $\dot{x}_0 = 1$, $\dot{y}_0 = p$

$$x = \left(1 + \frac{p^2}{2}\right)t - \frac{p^2}{4}\sin t, \quad y = p\sin t. \tag{1.9.1}$$

Note that this curve hits the line $y = 0$ in the point $x = \pi(1 + p^2/2)$.

We can compute the length of a geodesic by

$$\begin{aligned} L(\gamma) &:= \int_0^t \sqrt{m(\dot{x})^2 + \dot{y}} d\tau = \int_0^t \int \sqrt{\frac{C^2}{m} + D - \frac{C^2}{m}} d\tau \\ &= t\sqrt{D} = t\sqrt{m(y_0)(\dot{x}_0)^2 + (\dot{y}_0)^2}. \end{aligned}$$

We are thus ready to prove the first property of our space.

Lemma 1.67. *The minimal geodesics are not horizontal.*

Proof. From

$$\dot{y}^2 + \frac{C^2}{m(y)} = D$$

it follows that if $\dot{y} = 0$ then $m(y) = C^2/D$. Computing in those points

$$\frac{d^2y}{dx^2} = \frac{\ddot{y}}{(\dot{x})^2} - \frac{\dot{y}\ddot{x}}{(\dot{x})^3} = \frac{m'}{2} + \frac{m'}{m}\frac{(\dot{y})^2}{(\dot{x})^2} = \frac{m'}{2} \neq 0, \quad y \neq 0.$$

In ths other case $y = 0$, and thus the curve is $x = t$, $y = 0$. Computing the lenght of this curve starting from $x_0 = 0$ we obtain $L = x$, while if $|x| > \pi$ we can arrive from one of the curves (1.9.1) obtaining

$$L(p) = \pi\sqrt{1 + p^2} < \pi\left(1 + \frac{p^2}{2}\right) = |x|.$$

Hence also $\{y = 0\}$ is a geodesic only for a lenght 2π. $\square$

We can characterize more precisely the distance

$$\tilde{d}\big((x, y), (x', y')\big) := \inf\big\{L(\gamma), \gamma(0) = (x, y), \gamma(1) = (x', y')\big\}$$

on the line $\{y = 0\}$ as follows.

Lemma 1.68. *The distance function restricted on $y = 0$ is a concave function with the following properties:*

1. *it is translation invariant,*
2. $\tilde{d}(0, x) \geq |x|/4$,
3. $\tilde{d}(0, x) = |x|$ *for* $|x| \leq \pi$;
4. $\tilde{d}(0, x) = \sqrt{2\pi x - \pi^2}$ *for* $|x| \in (\pi, 3\pi/2)$.

Proof. The first two points are trivial since $M \geq \mathbb{I}/4$. For the second, observe that we can lower the distance by taking $m = (1 + y^2)^{-1}$, for which we can explicitly compute the distance as

$$\tilde{d}(x, 0) = \begin{cases} |x| & |x| \leq \pi, \\ \sqrt{2\pi x - \pi^2} & |x| > \pi. \end{cases}$$

Since for $p < 1$ the solutions remains in the strip $\{|y| \leq 1\}$, which corresponds to the solution (1.9.1) with $p = 1$, also the third point follows. $\square$

We now restrict the analysis to $y = 0$, with the distance d_L with the above properties. Consider the function ϕ_a, $a < 1$, defined up to a constant by

$$\phi'(x) = \begin{cases} 1 & x \in [z.z + a], \\ -1 & x \in (z + a, z + 1). \end{cases}$$

Lemma 1.69. *For $a < 5/9$ the function ϕ is a potential, i.e.*

$$\big|\phi(x') - \phi(x)\big| \leq \tilde{d}\big((x,0),(x',0)\big).$$

Proof. By translation and symmetry, the only case to study is $x = 0$, $x' > 0$. For $a < 5/9$ we have that at the maximum points

$$\phi(n+a) - \phi(0) = n(2a-1) + a = \frac{n}{9} + \frac{5}{9},$$

and thus

$$\phi(n+a) - \phi(0) \leq \begin{cases} n+a & n \leq 2, \\ (n+a)/4 & n \geq 3. \end{cases}$$

This concludes the proof. □

The key point of the above lemma is that in the geodesic space $(\mathbb{R}^2, \tilde{d})$ if μ, ν are two probability measures such that

$$\mu\big([n, n+5/18]\big) = \nu\big([n+5/18, n+5/18]\big),$$

then the usual monotone transport is $\tilde{d}$-monotone.

1.9.1 The space (X, d, d_L)

Consider the space

$$(X, d) := \big(\mathbb{T}^2 \times \mathbb{R}, |\cdot|\big), \quad w = (x,y,z),$$

and for $\alpha \in [0,1] \setminus \mathbb{Q}$ define

$$d_L\big((x,y,z),(x',y',x')\big) := \begin{cases} \tilde{d}\big((0,z),(t,z')\big) & x' - x = t \mod 1,\ z' - z = \alpha t \mod 1, \\ +\infty & \text{otherwise.} \end{cases}$$

Clearly the set $\{w' : d_L(w,w') < +\infty\}$ is the image of $(\mathbb{R}^2, \tilde{d})$ by the map

$$\mathbb{R}^2 \ni (t_1, t_2) \mapsto \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} x + t_1 \mod 1 \\ y + \alpha t_1 \mod 1 \\ z + t_2 \end{pmatrix}.$$

Since no geodesics are horizontal by Lemma 1.67, it follows that the local compactness condition is satisfied.

Consider now the (not renormalized) measures

$$\mu = \mathcal{H}^2 \llcorner_{[0,1/18]\times[0,1]\times\{0\}}, \quad \mu = \mathcal{H}^2 \llcorner_{[1/2,5/9]\times[0,1]\times\{0\}},$$

and the transport maps

$$T_1(x,y,z) := \left(x + \frac{1}{2}, y + \frac{\alpha}{2}, z\right), \quad (x,y,z) \in [0,1/18] \times [0,1] \times \{0\},$$

$$T_2 := \begin{cases} \big(x + \frac{17}{36}, y + \frac{17\alpha}{36}, z\big) & (x,y,z) \in [1/36, 1/18] \times [0,1] \times \{0\}, \\ \big(x - \frac{17}{36} \mod 1, y - \frac{17\alpha}{36} \mod 1, z\big) & (x,y,z) \in [0, 1/36) \times [0,1] \times \{0\}. \end{cases}$$

Clearly the conditions on the absolute continuity of the disintegration is satisfied, and the transport costs can be computed to be

$$\int d_L(x, T_1(x))\mu(dx) = \frac{1}{2}, \quad \int d_L(x, T_2(x))\mu(dx) = \frac{17}{36}.$$

Nevertheless T_1 is d_L monotone, because of Lemma 1.69, and the disintegration along the transport set is clearly a.c. w.r.t. $\mathcal{H}^1$.

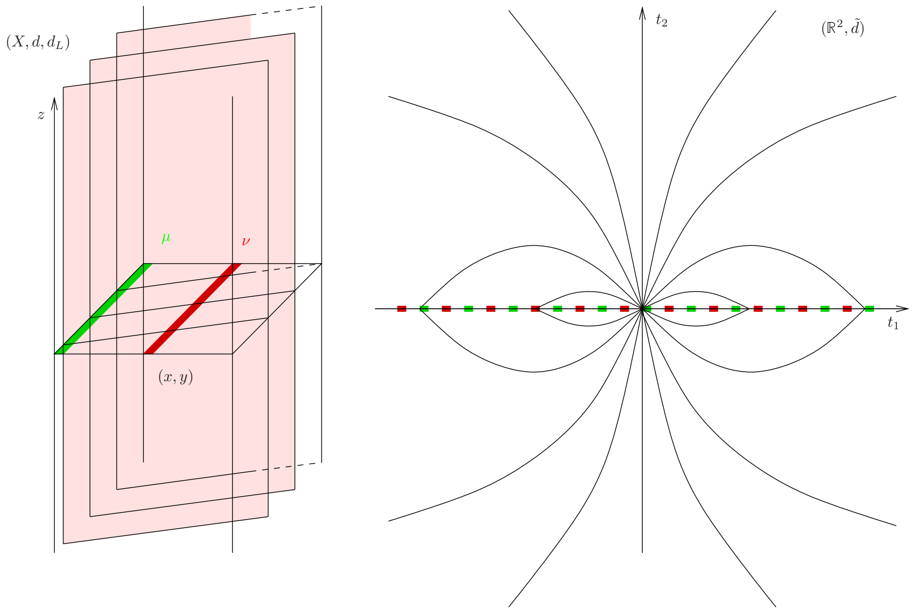

Figure 1.6: The metric space of Example.

Chapter 2

Optimal Transportation in Wiener space

We recall the issue of this chapter. Let $(X, \|\cdot\|)$ be an ∞-dimensional separable Banach space, $\gamma \in \mathcal{P}(X)$ be a non degenerate Gaussian measure over X and $H(\gamma)$ be the corresponding Cameron-Martin space with Hilbertian norm $\|\cdot\|_{H(\gamma)}$. Given two probability measures $\mu, \nu \in \mathcal{P}(X)$, we will prove the existence of a solution for the following Monge minimization problem

$$\min_{T:T_\sharp \mu=\nu} \int_X \|x - T(x)\|_{H(\gamma)} \mu(dx), \tag{2.0.1}$$

provided μ and ν are both absolutely continuous w.r.t. γ.

Just few words on the organization of this chapter. In Section 2.1 we recall the basic mathematical results we use: projective set theory, the Disintegration Theorem in the version of [8], selection principles, some fundamental results in optimal transportation theory and the definition and some properties of the abstract Wiener space.

In Section 2.2 we show, omitting the proof, the construction done in [9] on the Monge problem in a generalized non-branching geodesic space and we show that the Wiener space fits into the general setting.

In Section 2.3 we prove Theorem 0.4. In Section 2.4 we prove that the hypothesis of Theorem 0.4 can be proved by a finite-dimensional approximation and Section 2.5 proves the hypothesis of Theorem 0.4 in the finite dimensional case. Finally in Section 2.6 we prove Theorem 2.22 and we obtain the existence of an optimal transport map.

2.1 Preliminaries

2.1.1 Approximate differentiability of transport maps

The following results are taken from [2] where they are presented in full generality.

Definition 2.1 (Approximate limit and approximate differential)**.** Let $\Omega \subset \mathbb{R}^d$ be an open set and $f : \Omega \to \mathbb{R}^m$. We say that f has an approximate limit (respectively, approximate differential) at $x \in \Omega$ if there exists a function $g : \Omega \to \mathbb{R}^m$ continuous (resp. differentiable) at x such that the set $\{f \neq g\}$ has Lebesgue-density 0 at x. In this case the approximate limit (resp. approximate differential) will be denoted by $\tilde{f}(x)$ (resp. $\tilde{\nabla} f(x)$).

Recall that if $f : \Omega \to \mathbb{R}^m$ is $\mathcal{L}^d$-measurable, then it has approximate limit $\tilde{f}(x)$ at $\mathcal{L}^d$-a.e. $x \in \Omega$ and $f(x) = \tilde{f}(x)$ $\mathcal{L}^d$-a.e..

Consider $m = d$ and denote with Σ_f the Borel set of points where f is approximately differentiable.

This chapter is based on the work [16].

Lemma 2.2 (Density of the push-forward). *Let $\rho \in L^1(\mathbb{R}^d)$ be a nonnegative function and assume that there exists a Borel set $\Sigma \subset \Sigma_f$ such that $\tilde{f}\llcorner_\Sigma$ is injective and $\{\rho > 0\} \setminus \Sigma$ is $\mathcal{L}^d$-negligible. Then $f_\sharp \rho \mathcal{L}^d \ll \mathcal{L}^d$ if and only if $|\det \tilde{\nabla} f| > 0$ for $\mathcal{L}^d$-a.e. on Σ and in this case*

$$f_\sharp(\rho \mathcal{L}^d) = \frac{\rho}{|\det \tilde{\nabla} f|} \circ \tilde{f}^{-1} \llcorner_{f(\Sigma)} \mathcal{L}^d. \tag{2.1.1}$$

We include a regularity result for the Monge minimization problem in $\mathbb{R}^d$ with cost $c_p(x,y) = |x-y|^p$, $p > 1$ (Theorem 6.2.7 of [2]):

$$\min_{T : T_\sharp \mu = \nu} \int_{\mathbb{R}^d} c_p(x, T(x)) \mu(dx). \tag{2.1.2}$$

Theorem 2.3. *Assume that $\mu \in \mathcal{P}^r(\mathbb{R}^d)$, $\nu \in \mathcal{P}(\mathbb{R}^d)$,*

$$\mu\Big(\Big\{x \in \mathbb{R}^d : \int c_p(x,y)\nu(dy) < +\infty\Big\}\Big) > 0 \quad \text{and} \quad \nu\Big(\Big\{y \in \mathbb{R}^d : \int c_p(x,y)\mu(dx) < +\infty\Big\}\Big) > 0.$$

If the minimum of (5.4.1) *is finite, then*

i) there exists a unique solution T_p for the Monge problem (2.1.2)*;*

ii) for μ-a.e. $x \in \mathbb{R}^d$ the map T_p is approximately differentiable at x and $\tilde{\nabla} T_p(x)$ is diagonalizable with nonnegative eigenvalues.

2.1.2 The Abstract Wiener space

We briefly introduce our setting. The main reference is [11].

Given an infinite dimensional separable Banach space X, we denote by $\|\cdot\|_X$ its norm and X^* denotes the topological dual, with duality $\langle \cdot, \cdot \rangle$. Given the elements $x_1^*, \ldots, x_m^*$ in X^*, we denote by $\Pi_{x_1^*,\ldots,x_m^*} : X \to \mathbb{R}^m$ the map

$$\Pi_{x_1^*,\ldots,x_m^*}(x) := (\langle x, x_1^* \rangle, ..., \langle x, x_m^* \rangle).$$

Denoted with $\mathcal{E}(X)$ the σ-algebra generated by X^*. A set $C \in \mathcal{E}(X)$ is called *cylindrical* if

$$C = \{x \in X : \Pi_{\{x_i^*\}}(x) \in B\}, \quad B \subset \mathbb{R}^n, \quad \{x_i^*\}_{i \le n} \subset X^*,$$

and we will denote the cylindrical sets with $C(B)$ where B is the base of C.

A set E belongs to $\mathcal{E}(X)$ precisely when it has the form

$$E = \{x \in X : \Pi_{\{x_i^*\}}(x) \in B\}, \quad B \subset \mathbb{R}^\infty, \quad \{x_i^*\}_{i \in \mathbb{N}} \subset X^*.$$

In our setting $\mathcal{B}(X) = \mathcal{E}(X)$.

Lemma 2.4 (Lemma 2.1.5 of [11]). *Let μ be a positive Borel measure on X. For any set $A \in \mathcal{B}(X)_\mu$ (the completion of $\mathcal{B}(X)$ w.r.t. μ) and any $\varepsilon > 0$ there exists a set $E = C(B)$ in $\mathcal{E}(X)$ with $B \subset \mathbb{R}^\infty$ compact in the locally convex topology of $\mathbb{R}^\infty$, such that*

$$E \subset A, \qquad \mu(A \setminus E) < \varepsilon.$$

A Borel measure $\gamma \in \mathcal{P}(X)$ is a *non-degenerate centred Gaussian measure* if it is not concentrated on a proper closed subspace of X and for every $x^* \in X^*$ the measure $x^*_\sharp \gamma$ is a centred Gaussian measure on $\mathbb{R}$, that is, the Fourier transform of γ is given by

$$\hat{\gamma}(x^*) = \int_X \exp\{i\langle x^*, x\rangle\}\gamma(dx) = \exp\Big\{-\frac{1}{2}\langle x^*, Qx^*\rangle\Big\}$$

where $Q \in L(X^*, X)$ is the covariance operator. The non-degeneracy hypothesis of γ is equivalent to $\langle x^*, Qx^* \rangle > 0$ for every $x^* \neq 0$. The covariance operator Q is symmetric, positive and uniquely determined by the relation

$$\langle y^*, Qx^* \rangle = \int_X \langle x^*, x \rangle \langle y^*, x \rangle \gamma(dx), \quad \forall x^*, y^* \in X^*.$$

The fact that Q is bounded follows from the Fernique's Theorem, see [11]. This imply that any $x^* \in X^*$ defines a function $x \mapsto x^*(x)$ that belongs to $L^p(X, \gamma)$ for all $1 \leq p < \infty$. In particular let us denote by $R^*_\gamma : X^* \to L^2(X, \gamma)$ the embedding $R^*_\gamma x^*(x) := \langle x^*, x \rangle$. The space $\mathscr{H}$ given by the closure of $R^*_\gamma X^*$ in $L^2(X, \gamma)$ is called the *reproducing kernel* of the Gaussian measure. The definition is motivated by the fact that if we consider the operator $R_\gamma : \mathscr{H} \to X$ whose adjoint is R^*_γ then $Q = R_\gamma R^*_\gamma$:

$$\langle y^*, R_\gamma R^*_\gamma x^* \rangle = \langle R^*_\gamma y^*, R^*_\gamma x^* \rangle_{\mathscr{H}} = \int_X \langle x^*, x \rangle \langle y^*, x \rangle \gamma(dx) = \langle y^*, Qx^* \rangle.$$

It can proven that R_γ is injective, compact and that

$$R_\gamma \hat{h} = \int_X \hat{h}(x) x \gamma(dx), \quad \hat{h} \in \mathscr{H}, \tag{2.1.3}$$

where the integral is understood in the Bochner or Pettis sense.

The space $H(\gamma) = R_\gamma \mathscr{H} \subset X$ is called the Cameron-Martin space. It is a separable Hilbert space with inner product inherited from $L^2(X, \gamma)$ via R_γ:

$$\langle h_1, h_2 \rangle_{H(\gamma)} = \langle \hat{h}_1, \hat{h}_2 \rangle_{\mathscr{H}}.$$

for all $h_1, h_2 \in H$ with $h_i = R_\gamma \hat{h}_i$ for $i = 1, 2$. Moreover H is a dense subspace of X and by the compactness of R_γ follows that the embedding of $(H(\gamma), \| \cdot \|_{H(\gamma)})$ into $(X, \| \cdot \|)$ is compact. Note that if X is infinite dimensional then $\gamma(H) = 0$ and if X is finite dimensional then $X = H(\gamma)$.

2.1.3 Finite dimensional approximations

Using the embedding of X^* in $L^2(X, \gamma)$ we say that a family $\{x^*_i\} \subset X^*$ is orthonormal if the corresponding family $\{R^*_\gamma x^*_i\}$ is orthonormal in $\mathscr{H}$. In particular starting from a sequence $\{y^*_i\}_{i \in \mathbb{N}}$ whose image under R^*_γ is dense in $\mathscr{H}$, we can obtain an orthonormal basis $R^*_\gamma x^*_i$ of $\mathscr{H}$. Therefore also $h_j = R_\gamma R^*_\gamma x^*_j$ provide an orthonormal basis in $H(\gamma)$.

In the following we will consider a fixed orthonormal basis $\{e_i\}$ of $H(\gamma)$ with $e_i = R_\gamma \hat{e}_i$ for $\hat{e}_i \in R^*_\gamma X^*$.

Proposition 2.5 (Proposition 3.8.12 of [11])**.** *Let γ be a centred Gaussian measure on a Banach space X and $\{e_i\}$ an orthonormal basis in $H(\gamma)$. Define $P_d x := \sum_{i=1}^d \langle \hat{e}_i, x \rangle e_i$. Then the sequence of measures $\gamma_d := P_{d\,\sharp} \gamma \in \mathcal{P}(X)$ converges weakly to γ.*

The measure γ_d defined above is a centred non-degenerate d-dimensional Gaussian measure and, due to the orthonormality of $\{e_i\}_{i \in \mathbb{N}}$, with identity covariance matrix. Note that from (2.1.3) it follows that $\langle \hat{e}_j, x \rangle = \langle e_i, x \rangle_H$ for all $x \in H$. Hence we will not specify whether the measures γ_d is probability measures on $\mathbb{R}^d$ or on $P_d H$:

$$\gamma_d = \hat{e}_{1\,\sharp} \gamma \otimes \cdots \otimes \hat{e}_{d\,\sharp} \gamma, \qquad \hat{e}_{j\,\sharp} \gamma = \frac{1}{\sqrt{2\pi}} \exp\Big\{ -\frac{x^2}{2} \Big\} \mathcal{L}^1.$$

For every $d \in \mathbb{N}$ we can disintegrate γ w.r.t. the partition induced by the saturated sets of P_d:

$$\gamma = \int \gamma^\perp_{y,d} \gamma_d(dy), \quad \gamma^\perp_{y,d}(P_d^{-1}(y)) = 1 \quad \text{for } \gamma_d - \text{a.e. } y. \tag{2.1.4}$$

2.2 Optimal transportation in geodesic spaces

In what follows (X, d, d_L) is a generalized non-branching geodesic space in the sense of Chapter 1 and in this Section we retrace, omitting the proof, the construction done in Chapter 1 that permits to reduce the Monge problem with non-branching geodesic distance cost d_L, to a family of one dimensional transportation problems. Then we will observe that the triple $(X, \|\cdot\|, \|\cdot\|_{H(\gamma)})$ is a generalized non-branching geodesic space in the sense of Chapter 1.

Let $\mu, \nu \in \mathcal{P}(X)$ and let $\pi \in \Pi(\mu, \nu)$ be a d_L-cyclically monotone transference plan with finite cost. By inner regularity, we can assume that the optimal transference plan is concentrated on a σ-compact d_L-cyclically monotone set $\Gamma \subset \{d_L(x, y) < +\infty\}$. By Lusin Theorem, we can require also that $d_L \llcorner_\Gamma$ is σ-continuous:

$$\Gamma = \cup_n \Gamma_n, \ \Gamma_n \subset \Gamma_{n+1} \text{ compact}, \quad d_L \llcorner_{\Gamma_n} \text{ continuous}.$$

Consider the set

$$\begin{aligned}\Gamma' := \Big\{ &(x, y) : \exists I \in \mathbb{N}_0, (w_i, z_i) \in \Gamma \text{ for } i = 0, \dots, I, \ z_I = y \\ &w_{I+1} = w_0 = x, \ \sum_{i=0}^{I} d_L(w_{i+1}, z_i) - d_L(w_i, z_i) = 0 \Big\}. \end{aligned} \tag{2.2.1}$$

In other words, we concatenate points $(x, z), (w, y) \in \Gamma$ if they are initial and final point of a cycle with total cost 0. One can prove that $\Gamma \subset \Gamma' \subset \{d_L(x, y) < +\infty\}$, if Γ is analytic so is Γ' and if Γ is d_L-cyclically monotone so is Γ'.

Definition 2.6. [Transport rays] Define the *set of oriented transport rays*

$$G := \Big\{ (x, y) : \exists (w, z) \in \Gamma', d_L(w, x) + d_L(x, y) + d_L(y, z) = d_L(w, z) \Big\}.$$

For $x \in X$, the *outgoing transport rays from* x is the set $G(x)$ and the *incoming transport rays in* x is the set $G^{-1}(x)$. Define the *set of transport rays* as the set

$$R := G \cup G^{-1}.$$

It is fairly easy to prove that G is still d_L-cyclically monotone, $\Gamma' \subset G \subset \{d_L(x, y) < +\infty\}$ and G and R are analytic sets.

Definition 2.7. Define the *transport sets*

$$\begin{aligned}\mathcal{T} &:= P_1\big(\mathrm{graph}(G^{-1}) \setminus \{x = y\}\big) \cap P_1\big(\mathrm{graph}(G) \setminus \{x = y\}\big), \\ \mathcal{T}_e &:= P_1\big(\mathrm{graph}(G^{-1}) \setminus \{x = y\}\big) \cup P_1\big(\mathrm{graph}(G) \setminus \{x = y\}\big).\end{aligned}$$

From the definition of G one can prove that $\mathcal{T}$, $\mathcal{T}_e$ are analytic sets. The subscript e refers to the endpoints of the geodesics: we have

$$\mathcal{T}_e = P_1(R \setminus \{x = y\}). \tag{2.2.2}$$

It follows that we have only to study the Monge problem in $\mathcal{T}_e$: $\pi(\mathcal{T}_e \times \mathcal{T}_e \cup \{x = y\}) = 1$. As a consequence, $\mu(\mathcal{T}_e) = \nu(\mathcal{T}_e)$ and any maps T such that for $\nu \llcorner_{\mathcal{T}_e} = T_\sharp \mu \llcorner_{\mathcal{T}_e}$ can be extended to a map T' such that $\nu = T_\sharp \mu$ with the same cost by setting

$$T'(x) = \begin{cases} T(x) & x \in \mathcal{T}_e \\ x & x \notin \mathcal{T}_e. \end{cases} \tag{2.2.3}$$

By the non-branching assumption, if $x \in \mathcal{T}$, then $R(x)$ is a single geodesic and therefore the set $R \cap \mathcal{T} \times \mathcal{T}$ is an equivalence relation on $\mathcal{T}$ that we will call ray equivalence relation. Notice that the set G is a partial order relation on $\mathcal{T}_e$.

The next step is to study the set $\mathcal{T}_e \setminus \mathcal{T}$.

Definition 2.8. Define the multivalued *endpoint graphs* by:

$$a := \{(x,y) \in G^{-1} : G^{-1}(y) \setminus \{y\} = \emptyset\},$$
$$b := \{(x,y) \in G : G(y) \setminus \{y\} = \emptyset\}.$$

We call $P_2(a)$ the set of *initial points* and $P_2(b)$ the set of *final points.*

Even if a, b are not in the analytic class, still they belong to the σ-algebra $\mathcal{A}$.

Proposition 2.9. *The following holds:*

1. *the sets*
$$a, b \subset X \times X, \quad a(A), b(A) \subset X,$$
belong to the $\mathcal{A}$-class if A analytic;
2. $a \cap b \cap \mathcal{T}_e \times X = \emptyset$;
3. $a(x)$, $b(x)$ *are singleton or empty when* $x \in \mathcal{T}$;
4. $a(\mathcal{T}) = a(\mathcal{T}_e)$, $b(\mathcal{T}) = b(\mathcal{T}_e)$;
5. $\mathcal{T}_e = \mathcal{T} \cup a(\mathcal{T}) \cup b(\mathcal{T})$, $\mathcal{T} \cap (a(\mathcal{T}) \cup b(\mathcal{T})) = \emptyset$.

Finally we can assume that the μ-measure of final points and the ν-measure of the initial points are 0: indeed since the sets $G \cap b(\mathcal{T}) \times X$, $G \cap X \times a(\mathcal{T})$ is a subset of the graph of the identity map, it follows that from the definition of b one has that

$$x \in b(\mathcal{T}) \quad \Longrightarrow \quad G(x) \setminus \{x\} = \emptyset,$$

A similar computation holds for a. Hence we conclude that

$$\pi(b(\mathcal{T}) \times X) = \pi(G \cap b(\mathcal{T}) \times X) = \pi(\{x = y\}),$$

and following (2.2.3) we can assume that

$$\mu(b(\mathcal{T})) = \nu(a(\mathcal{T})) = 0.$$

2.2.1 The Wiener case

For the abstract Wiener space, it is possible to obtain more regularity for the sets introduced so far. Let $d = \|\cdot\|$ and $d_L = \|\cdot\|_H$: by the compactness of the embedding R_γ of H into X it follows that

(1) $d_L : X \times X \to [0, +\infty]$ l.s.c. distance;

(2) $d_L(x,y) \geq Cd(x,y)$ for some positive constant C;

(3) $\cup_{x \in K_1, y \in K_2} \gamma_{[x,y]}$ is d-compact if K_1, K_2 are d-compact, $d_{L \llcorner K_1 \times K_2}$ uniformly bounded.

The set Γ' is σ-compact: in fact, if one restrict to each Γ_n given by (2.2.1), then the set of cycles of order I is compact, and thus

$$\Gamma'_{n,\bar{I}} := \Bigg\{(x,y) : \exists I \in \{0, \dots, \bar{I}\}, (w_i, z_i) \in \Gamma_n \text{ for } i = 0, \dots, I, \ z_I = y$$
$$w_{I+1} = w_0 = x, \ \sum_{i=0}^{I} d_L(w_{i+1}, z_i) - d_L(w_i, z_i) = 0\Bigg\}$$

is compact. Finally $\Gamma' = \cup_{n,I} \Gamma'_{n,I}$.

Moreover, $d_{L \llcorner \Gamma'_{n,I}}$ is continuous. If $(x_n, y_n) \to (x, y)$, then from the l.s.c. and

$$\sum_{i=0}^{I} d_L(w_{n,i+1}, z_{n,i}) = \sum_{i=0}^{I} d_L(w_{n,i}, z_{n,i}), \quad w_{n,I+1} = w_{n,0} = x_n, \; z_{n,I} = y_n,$$

it follows also that each $d_L(w_{n,i+1}, z_{n,i})$ is continuous.

Similarly the sets G, R, a, b are σ-compact: assumption (3) and the above computation in fact shows that

$$G_{n,I} := \Big\{(x,y) : \exists (w,z) \in \Gamma'_{n,I}, d_L(w,x) + d_L(x,y) + d_L(y,z) = d_L(w,z)\Big\}$$

is compact. For a, b, one uses the fact that projection of σ-compact sets is σ-compact.

So we have that Γ, Γ', G, G^{-1}, a and b are σ-compact sets.

2.2.2 Strongly consistency of disintegrations

The strong consistency of the disintegration follows from the next result.

Proposition 2.10. *There exists a μ-measurable cross section $f : \mathcal{T} \to \mathcal{T}$ for the ray equivalence relation R.*

Up to a μ-negligible saturated set $\mathcal{T}_N$, we can assume it to have σ-compact range: just let $S \subset f(\mathcal{T})$ be a σ-compact set where $f_\sharp \mu \llcorner_{\mathcal{T}}$ is concentrated, and set

$$\mathcal{T}_S := R^{-1}(S) \cap \mathcal{T}, \quad \mathcal{T}_N := \mathcal{T} \setminus \mathcal{T}_S, \quad \mu(\mathcal{T}_N) = 0. \tag{2.2.4}$$

Having the $\mu \llcorner_{\mathcal{T}}$-measurable cross-section

$$\mathcal{S} := f(\mathcal{T}) = S \cup f(\mathcal{T}_N) = (\text{Borel}) \cup (f(\mu\text{-negligible})),$$

we can define the parametrization of $\mathcal{T}$ and $\mathcal{T}_e$ by geodesics.

Using the quotient map f, we obtain a unitary speed parametrization of the transport set.

Definition 2.11 (Ray map)**.** Define the *ray map* g by the formula

$$\begin{aligned} g := & \Big\{(y,t,x) : y \in \mathcal{S}, t \in [0,+\infty), x \in G(y) \cap \{d_L(x,y) = t\}\Big\} \\ & \cup \Big\{(y,t,x) : y \in \mathcal{S}, t \in (-\infty, 0), x \in G^{-1}(y) \cap \{d_L(x,y) = -t\}\Big\} \\ = & \, g^+ \cup g^-. \end{aligned}$$

Proposition 2.12. *The following holds.*

1. *The restriction $g \cap S \times \mathbb{R} \times X$ is analytic.*
2. *The set g is the graph of a map with range $\mathcal{T}_e$.*
3. *$t \mapsto g(y,t)$ is a d_L 1-Lipschitz G-order preserving for $y \in \mathcal{T}$.*
4. *$(t,y) \mapsto g(y,t)$ is bijective on $\mathcal{T}$, and its inverse is*

$$x \mapsto g^{-1}(x) = \big(f(y), \pm d_L(x, f(y))\big)$$

where f is the quotient map of Proposition 2.10 and the positive/negative sign depends on $x \in G(f(y))/x \in G^{-1}(f(y))$.

Another property of d_L-cyclically monotone transference plans.

Proposition 2.13. *For any π d_L-monotone there exists a d_L-cyclically monotone transference plan $\tilde{\pi}$ with the same cost of π such that it coincides with the identity on $\mu \wedge \nu$.*

Coming back to the abstract Wiener space, we have that given $\mu, \nu \ll \gamma$ and given $\pi \in \Pi(\mu,\nu)$ $\|\cdot\|_{H(\gamma)}$-cyclically monotone, we have constructed the transport $\mathcal{T}$ (and $\mathcal{T}_e$), an equivalence relation R on it with geodesics as equivalence classes and the corresponding disintegration is strongly consistent:

$$\mu\llcorner_{\mathcal{T}} = \int_{\mathcal{S}} \mu_y m(dy) \tag{2.2.5}$$

with $m = f_\sharp \mu$ and $\mu_y(R(y)) = 1$ for m-a.e. $y \in \mathcal{T}$. Using the ray map g one can assume that $\mu_y \in \mathcal{P}(\mathbb{R})$ and

$$\mu\llcorner_{\mathcal{T}} = g_\sharp \int_{\mathcal{S}} \mu_y m(dy).$$

2.3 Regularity of disintegration

To obtain existence of an optimal transport map it is enough to prove that:

- μ in concentrated on $\mathcal{T}$;
- μ_y is a continuous measure for m-a.e. $y \in \mathcal{S}$.

Indeed at that point, for every $y \in \mathcal{S}$ we consider the unique monotone map T_y such that $T_{y\,\sharp}\mu_y = \nu_y$, then $T(g(y,t)) := T_y(g(y,t))$ is an optimal transport map, see Theorem 1.37 of Chapter 1.

Define the map $X \times X \ni (x,y) \mapsto T_t(x,y) := x(1-t) + yt \in X$.

Assumption 5 (Non-degeneracy assumption)**.** The measure γ is said to satisfy Assumption 5 w.r.t. a $\|\cdot\|_{H(\gamma)}$-cyclically monotone set Γ if

i) $\pi(\Gamma) = 1$ with $\pi \in \Pi(\mu,\nu)$ and $\mu, \nu \ll \gamma$;

ii) for each set $A \in \mathcal{E}(X)$ with compact base such that $\mu(A) > 0$ there exist $C > 0$ and $\{t_n\}_{n\in\mathbb{N}} \subset [0,1]$ converging to 0 as $n \to +\infty$ such that

$$\gamma(T_{t_n}(\Gamma \cap A \times X)) \geq C\mu(A) \tag{2.3.1}$$

for all $n \in \mathbb{N}$.

An immediate consequence of Assumption 5 is that the set of final points is γ-negligible.

Proposition 2.14. *If γ satisfies Assumption* (5) *then*

$$\mu(a(\mathcal{T}_e)) = 0.$$

Proof. Let $A = a(\mathcal{T}_e)$ and recall that $\mu = \rho_1\gamma$. Suppose by contradiction $\mu(A) > 0$. By inner regularity and Lemma 2.4 there exist a Borel set $C(B) =: \hat{A} \subset A$, with compact base B, of positive μ-measure and a strictly positive constant $\delta \in \mathbb{R}$ such that $\rho_1(x) \geq \delta$ for all $x \in \hat{A}$. Since $\Gamma \subset \{(x,y) : \|x-y\|_{H(\gamma)} < +\infty\}$, we can moreover assume that

$$\Gamma \cap \hat{A} \times X \subset \{(x,y) : \|x-y\|_{H(\gamma)} \leq M\}$$

for some positive $M \in \mathbb{R}$.

By Assumption 5 there exist $C > 0$ and $\{t_n\}_{n\in\mathbb{N}}$ converging to 0 such that

$$\gamma(T_{t_n}(\Gamma \cap \hat{A} \times X)) \geq C\mu(\hat{A}) \geq \delta C\gamma(\hat{A}).$$

Denote with $\hat{A}_{t_n} = T_{t_n}(\Gamma \cap \hat{A} \times X)$ and define

$$\hat{A}^\varepsilon := \big\{x : \|\hat{A} - x\|_{H(\gamma)} < \varepsilon\big\} = P_1\Big(\big\{(x, y) \in X \times \hat{A} : \|x - y\|_{H(\gamma)} < \varepsilon\big\}\Big).$$

Since $\hat{A} \subset A = a(\mathcal{T}_e)$, $\hat{A}_{t_n} \cap \hat{A} = \emptyset$ for every $n \in \mathbb{N}$. Moreover for $t_n \leq \varepsilon/M$ it holds $\hat{A}^\varepsilon \supset \hat{A}_{t_n}$. So we have for t_n small enough

$$\gamma(\hat{A}^\varepsilon) \geq \gamma(\hat{A}) + \gamma(\hat{A}_{t_n}) \geq (1 + C\delta)\gamma(\hat{A}).$$

Since $\gamma(\hat{A}) = \lim_{\varepsilon\to 0}\gamma(\hat{A}^\varepsilon)$, this is a contradiction. □

It follows that $\mu(\mathcal{T}) = 1$, therefore we can use the Disintegration Theorem 5.3 to write

$$\mu = \int_S \mu_y m(dy), \quad m = f_\sharp\mu, \ \mu_y \in \mathcal{P}(R(y)). \tag{2.3.2}$$

The disintegration is strongly consistent since the quotient map $f : \mathcal{T} \to \mathcal{T}$ is μ-measurable and $(\mathcal{T}, \mathcal{B}(\mathcal{T}))$ is countably generated.

The second consequence of Assumption 5 is that μ_y is continuous, i.e. $\mu_y(\{x\}) = 0$ for all $x \in X$.

Proposition 2.15. *If γ satisfies Assumption 5 then the conditional probabilities μ_y are continuous for m_γ-a.e. $y \in S$.*

Proof. From the regularity of the disintegration and the fact that $m(S) = 1$, we can assume that the map $y \mapsto \mu_y$ is weakly continuous on a compact set $K \subset S$ of comeasure $< \varepsilon$. It is enough to prove the proposition on K.

Step 1. From the continuity of $K \ni y \mapsto \mu_y \in \mathcal{P}(X)$ w.r.t. the weak topology, it follows that the map

$$y \mapsto A(y) := \big\{x \in R(y) : \mu_y(\{x\}) > 0\big\} = \cup_n\big\{x \in R(y) : \mu_y(\{x\}) \geq 2^{-n}\big\}$$

is σ-closed: in fact, if $(y_m, x_m) \to (y, x)$ and $\mu_{y_m}(\{x_m\}) \geq 2^{-n}$, then $\mu_y(\{x\}) \geq 2^{-n}$ by u.s.c. on compact sets. Hence A is Borel.

Step 2. The claim is equivalent to $\mu(P_2(A)) = 0$. Suppose by contradiction $\mu(P_2(A)) > 0$. By Lusin Theorem (Theorem 5.8.11 of [25]) A is the countable union of Borel graphs. Therefore we can take a Borel selection of A just considering one of the Borel graphs, say $\hat{A}$. Clearly $m(P_1(\hat{A})) > 0$ hence by (3.3.2) $\mu(P_2(\hat{A})) > 0$. Using Lemma 2.4 we can find a Borel subset $\tilde{A} \subset P_2(\hat{A})$ still with positive μ-measure such that $\tilde{A} = C(B)$ with $B \subset \mathbb{R}^\infty$ compact.

By Assumption 5, $\gamma(T_{t_n}(\Gamma\cap\tilde{A}\times X)) \geq C\mu(\tilde{A})$ for some $C > 0$ and $t_n \to 0$. From $T_{t_n}(\Gamma\cap\tilde{A}\times X)\cap(\tilde{A}) = \emptyset$, using the same argument of Proposition 2.14, the claim follows. □

2.4 An approximation result

Let $P_d : X \to H$ be the projection map of Proposition 2.5 associated to the orthonormal basis $\{e_i\}_{i\in\mathbb{N}}$ of $H(\gamma)$ with $e_i = R_\gamma \hat{e}_i$ for $\hat{e}_i \in R^*_\gamma X^*$ and $P_{d\,\sharp}\gamma = \gamma_d$.

Consider the following measures

$$\mu_d := P_{d\,\sharp}\mu, \qquad \nu_d := P_{d\,\sharp}\nu \tag{2.4.1}$$

and observe that $\mu_d = \rho_{1,d}\gamma_d$ and $\nu_d = \rho_{2,d}\gamma_d$ with

$$\rho_{i,d}(z) = \int \rho_i(x)\gamma^\perp_{z,d}(dx), \qquad i = 1, 2, \tag{2.4.2}$$

where $\gamma^\perp_{z,d}$ is defined in (2.1.4). Recall that $\mu_d \rightharpoonup \mu$ and $\nu_d \rightharpoonup \nu$ as $d \nearrow \infty$.

Denote with $\Pi_o(\mu_d, \nu_d)$ the set of optimal transference plans for the Monge problem between μ_d and ν_d with $\|\cdot\|_{H(\gamma)}$-cost.

Proposition 2.16. *Let $\pi_d \in \Pi_o(\mu_d, \nu_d)$ and let $\pi \in \Pi(\mu, \nu)$ be any weak limit of $\{\pi_d\}_{d\in\mathbb{N}}$. Then $\pi \in \Pi(\mu, \nu)$ is an optimal transport plan for* (2.0.1).

Proof. Let $\hat{\pi} \in \Pi(\mu, \nu)$ be a transference plan. The following holds true

$$\begin{aligned}\int \|x-y\|_H \hat{\pi}(dxdy) &\geq \int \|P_d(x-y)\|_{H(\gamma)}\hat{\pi}(dxdy) = \int \|x-y\|_{H(\gamma)}((P_d\otimes P_d)_\sharp \hat{\pi})(dxdy) \\ &\geq \int \|x-y\|_{H(\gamma)}\pi_d(dxdy).\end{aligned}$$

Let $\{d_k\}_{k\in\mathbb{N}}$ be a subsequence such that $\pi_{d_k} \rightharpoonup \pi$ as $d_k \nearrow \infty$. Since $\|\cdot\|_{H(\gamma)}$ is l.s.c. it follows that

$$\int \|x-y\|_{H(\gamma)}\hat{\pi}(dxdy) \geq \liminf_{k\to+\infty} \int \|x-y\|_{H(\gamma)}\pi_{d_k}(dxdy) \geq \int \|x-y\|_{H(\gamma)}\pi(dxdy).$$

Hence the claim follows. ☐

Since $\rho_{i,d}$ depend only on the first d-coordinates, the measures μ_d, ν_d can be considered also as probability measure on $\mathbb{R}^d$. Clearly for $x, y \in P_d(X)$ the norm $\|\cdot\|_d$ and $\|\cdot\|_{H(\gamma)}$ coincide. Therefore we can study the transport problem with euclidean norm cost $\|x\|_d^2 := \sum_{j=1}^d x_j^2$:

$$\min_{\pi\in\Pi(\mu_d,\nu_d)} \int \|x-y\|_d \pi(dxdy). \tag{2.4.3}$$

However it is worth noting that when we speak of weak convergence, the measures μ_d, ν_d and γ_d are all thought as probability measures in X.

It is a well-known fact in optimal transportation that (2.4.3) has a minimizer of the form $(Id, T_d)_\sharp\mu_d$ with T_d μ-essentially invertible and Borel. For each d we choose as optimal map T_d the one obtained gluing the monotone rearrangements over the geodesics and we set $\pi_d := (Id, T_d)_\sharp\mu_d$. Moreover $\Gamma_d := \mathrm{graph}(T_d)$.

The results that we are about to present are true for any weak limit π of the sequence $\{\pi_d\}_{d\in\mathbb{N}}$. Nevertheless to simplify the notation we assume that the whole sequence $\{\pi_d\}_{d\in\mathbb{N}} = \{(Id, T_d)_\sharp\mu_d\}_{d\in\mathbb{N}}$ converges to some π.

Theorem 2.17. *Fix $t \in [0, 1]$. Assume that there exists $C > 0$ such that for all $d \in \mathbb{N}$ and $A \subset X$ compact set the following holds true*

$$\gamma_d\big(T_t(\Gamma_d \cap A \times X)\big) \geq C\mu_d(A).$$

Then for all $A \subset X$ γ-measurable

$$\gamma\big(T_t(\Gamma \cap A \times X)\big) \geq C\mu(A), \tag{2.4.4}$$

where $\Gamma \subset X \times X$ is any $\|\cdot\|_{H(\gamma)}$-cyclically monotone with $\pi(\Gamma) = 1$.

Proof. It follows from Proposition 2.16 that π is an optimal transference plan, hence it is concentrated on a $\|\cdot\|_{H(\gamma)}$-cyclically monotone set Γ.

Step 1. Since $\mu_d \rightharpoonup \mu$ and $\nu_d \rightharpoonup \nu$, for every $\varepsilon > 0$ there exist $K_{1,\varepsilon}$ and $K_{2,\varepsilon}$ compact sets such that $\mu_d(K_{1,\varepsilon}) \geq 1 - \varepsilon/3$ and $\nu_d(K_{2,\varepsilon}) \geq 1 - \varepsilon/3$. Denote with $K_\varepsilon := K_{1,\varepsilon} \times K_{2,\varepsilon}$. For every $d \in \mathbb{N}$ there exists a compact set $\hat{\Gamma}_d \subset \Gamma_d$ such that $\pi_d(\hat{\Gamma}_d) \geq 1 - \varepsilon/3$. Consider the compact se $\Gamma_{d,\varepsilon} := \hat{\Gamma}_d \cap K_\varepsilon$, then $\pi_d(\Gamma_{d,\varepsilon}) \geq 1 - \varepsilon$ and $\Gamma_{d,\varepsilon}$ converges as $d \nearrow \infty$ in the Hausdorff topology, up to subsequences, to a compact set Γ_ε with $\pi(\Gamma_\varepsilon) \geq 1 - \varepsilon$.

Step 2. Let $\Gamma_n \subset \Gamma$ be a compact set such that $\pi(\Gamma_n) \geq 1 - 1/n$. Hence $\pi(\Gamma_\varepsilon \cap \Gamma_n) \geq 1 - \varepsilon - 1/n$. Consider the following sets, open and closed respectively:

$$(\Gamma_\varepsilon \cap \Gamma_n)^\delta := \{x : \|\Gamma_\varepsilon \cap \Gamma_n - x\| < \delta\}, \quad cl(\Gamma_\varepsilon \cap \Gamma_n)^\delta := \{x : \|\Gamma_\varepsilon \cap \Gamma_n - x\| \leq \delta\}.$$

Since $\liminf_d \pi_d(U) \geq \pi(U)$ for every open set $U \subset X$, it follows that for every $\delta > 0$ there exists $d_\delta \in \mathbb{N}$ such that for all $d \geq d_\delta$

$$\pi_d\big((\Gamma_\varepsilon \cap \Gamma_n)^\delta\big) \geq 1 - 2\varepsilon - 1/n.$$

The same inequality holds true for $\pi_d\big(cl(\Gamma_\varepsilon \cap \Gamma_n)^\delta\big)$. Therefore

$$\pi_{d_\delta}\big(\Gamma_{d_\delta,\varepsilon} \cap cl(\Gamma_\varepsilon \cap \Gamma_n)^\delta\big) \geq 1 - 3\varepsilon - 1/n.$$

Take as $\delta = 1/k$ for $k \in \mathbb{N}$ and let $d_k := d_{\delta_k}$. Define the compact set $\Gamma^n_{k,\varepsilon} := \Gamma_{d_k,\varepsilon} \cap cl(\Gamma_\varepsilon \cap \Gamma_n)^{1/k}$, then since $\Gamma^n_{k,\varepsilon} \subset K_\varepsilon$, up to subsequences, $\lim_{k\nearrow\infty} d_H(\Gamma^n_{k,\varepsilon}, \Gamma_{\varepsilon,n}) = 0$ with

$$\Gamma_{\varepsilon,n} \subset \Gamma_\varepsilon \cap \Gamma_n \subset \Gamma, \quad \pi_{d_k}(\Gamma^n_{k,\varepsilon}) \geq 1 - 3\varepsilon - 1/n, \quad \pi(\Gamma_{\varepsilon,n}) \geq 1 - 3\varepsilon - 1/n. \tag{2.4.5}$$

The inclusion $\Gamma_{\varepsilon,n} \subset \Gamma_\varepsilon \cap \Gamma_n$ can be verified observing that any limit point of sequences of $\Gamma^n_{k,\varepsilon}$ must be contained in $\Gamma_\varepsilon \cap \Gamma_n$.

Step 3. Consider $A = C(B) \in \mathcal{E}(X)$ with $B \in \mathbb{R}^m$ compact set for some fixed $m \in \mathbb{N}$. Since T_t is continuous and $\Gamma^n_{k,\varepsilon} \cap A \times X$ converges in Hausdorff topology to $\Gamma_{\varepsilon,n} \cap A \times X$, it is fairly easy to prove that $T_t(\Gamma^n_{k,\varepsilon} \cap A \times X)$ Kuratowski-converges to $T_t(\Gamma_{\varepsilon,n} \cap A \times X)$. For the definition of Kuratowski-convergence see for instance [6]. Moreover

$$T_t(\Gamma^n_{k,\varepsilon} \cap A \times X) \subset \overline{co}(P_1(K_\varepsilon \cap A \times X) \cap P_2(K_\varepsilon \cap A \times X))$$

and by Proposition A.1.6 of [11], $\overline{co}(P_1(K_\varepsilon \cap A \times X) \cap P_2(K_\varepsilon \cap A \times X))$ is compact. Therefore, by Proposition 4.4.14 of [6], $T_t(\Gamma^n_{k,\varepsilon} \cap A \times X)$ converges also in the Hausdorff topology to $T_t(\Gamma_{\varepsilon,n} \cap A \times X)$.

Step 4. It follows that

$$\gamma(T_t(\Gamma_{\varepsilon,n} \cap A \times X)) \geq \limsup_{k\to+\infty} \gamma_{d_k}(T_t(\Gamma^n_{k,\varepsilon} \cap A \times X)),$$

hence, since $\Gamma^n_{k,\varepsilon}$ is a subset of the graph Γ_{d_k}, it follows that

$$\begin{aligned}
\gamma(T_t(\Gamma_{\varepsilon,n} \cap A \times X)) &\geq \limsup_{k\to+\infty} \gamma_{d_k}(T_t(\Gamma^n_{k,\varepsilon} \cap A \times X)) \\
&\geq C \limsup_{k\to+\infty} \mu_{d_k}(P_1(\Gamma^n_{k,\varepsilon}) \cap A) \\
&\geq C \limsup_{k\to+\infty} \mu_{d_k}(A) - C(3\varepsilon - 1/n)
\end{aligned} \tag{2.4.6}$$

where in the last equation we have used $\mu_d(P_1(\Gamma^n_{k,\varepsilon})) \geq 1 - 3\varepsilon - 1/n$ that follows from (2.4.5). Since $\mu_{d_k} = P_{d_k\sharp}\mu$ and A has finite dimensional base, the sequence $\{\mu_{d_k}(A)\}_{k\in\mathbb{N}}$ is definitively constant and therefore

$$\gamma(T_t(\Gamma_{\varepsilon,n} \cap A \times X)) \geq C\mu(A) - C(3\varepsilon - 1/n) \tag{2.4.7}$$

for all $A \in \mathcal{E}(X)$ with finite dimensional base.

Step 5. Consider $A = C(B) = \{x \in X : \{\ell_i(x)\}_{i\in\mathbb{N}} \in B\}$ with $B \in \mathbb{R}^\infty$ compact set in the locally convex topology of $\mathbb{R}^\infty$ and $\{\ell_i\}_{i\in\mathbb{N}} \subset X^*$. We consider the sequence of compact sets

$$A_d := C(P_d(B)) = \{x \in X : \{\ell_i(x)\}_{i\leq d} \in P_d(B)\}.$$

Clearly A_d is closed with finite-dimensional compact base and $A_d \supset A_{d+1} \supset A$ for every $d \in \mathbb{N}$. Then for every $d \in \mathbb{N}$ from (2.4.7)

$$\gamma(T_t(\Gamma_{\varepsilon,n} \cap A_d \times X)) \geq C\mu(A_d) - C(3\varepsilon - 1/n) \geq C\mu(A) - C(3\varepsilon - 1/n).$$

Since the first term in the above inequality is decreasing, it follows that

$$\lim_{d\to+\infty} \gamma(T_t(\Gamma_{\varepsilon,n} \cap A_d \times X)) \geq C\mu(A) - C(3\varepsilon - 1/n).$$

Now observe that from the compactness of $\Gamma_{\varepsilon,n} \cap A_d \times X$ we obtain

$$\bigcap_{d=1}^{\infty} T_t(\Gamma_{\varepsilon,n} \cap A_d \times X) = T_t(\Gamma_{\varepsilon,n} \cap A \times X).$$

Thus (2.4.7) holds true and

$$\gamma(T_t(\Gamma \cap A \times X)) \geq \gamma(T_t(\Gamma_{\varepsilon,n} \cap A \times X)) \geq C\mu(A) - C(3\varepsilon - 1/n).$$

Letting $\varepsilon \to 0$ and $n \to +\infty$, the claim is proved for every $A \in \mathcal{E}(X)$ with compact base. The extension to γ-measurable sets is now a straightforward application of Lemma 2.4. $\square$

2.5 Finite dimensional estimate

The next theorem proves that the d-dimensional standard Gaussian measure $\gamma_d = P_{d\sharp}\gamma$ satisfies Assumption 5 for $\Gamma = \mathrm{graph}(T_d) = \Gamma_d$.

Theorem 2.18. *Assume that there exists $C > 0$ such that $\rho_{i,d}(x) \leq C$ for γ_d-a.e. $x \in \mathbb{R}^d$ and $i = 1, 2$. Then the following estimate holds true*

$$\gamma_d\big(T_t(\Gamma_d \cap A \times X)\big) \geq \frac{1}{C}\mu_d(A), \quad \forall t \in [0,1], A \in \mathcal{B}(\mathbb{R}^d).$$

Proof. Observe that the set $T_t(\Gamma_d \cap A \times X)$ is parametrized by the map $T_{d,t} := (1-t)Id + tT_d$.

Step 1. Consider the Monge minimization problem with cost c_p, (2.1.2), between μ_d and ν_d. It follows from Theorem 2.3 and from the boundedness of $\rho_{i,d}$ that there exists a unique optimal map $T_{p,d}$ approximately differentiable μ_d-a.e.. We will use the following notations: $\rho_i = \rho_{i,d}$ and $T_p = T_{p,d}$. By Lemma 2.2 it follows that

$$\rho_2(T_p(x))|\det \tilde{\nabla} T_p|(x) \prod_{j=1}^{d} \frac{1}{\sqrt{2\pi}} \exp\Big\{ - \frac{T_p(x)_j^2}{2} \Big\} = \rho_1(x) \prod_{j=1}^{d} \frac{1}{\sqrt{2\pi}} \exp\Big\{ - \frac{x_j^2}{2} \Big\}.$$

Since for μ_d-a.e. $x \in \mathbb{R}^d$ $|\det \tilde{\nabla} T_p|(x) > 0$, also $\rho_2(T_p(x)) > 0$ for μ_d-a.e. $x \in \mathbb{R}^d$. Hence the following makes sense μ-a.e.:

$$Jac(T_p)(x) = |\det \tilde{\nabla} T_p|(x) = \frac{\rho_1(x)}{\rho_2(T_p(x))} \exp\Big\{ \sum_{j=1}^{d} -\frac{1}{2}(x_j^2 - T_p(x)_j^2) \Big\}.$$

Step 2. Let $T_{p,t} := (1-t)Id + tT_p$. From Theorem 2.3, $\det \tilde{\nabla} T_p(x) = \prod_{j=1}^{d} \lambda_j$ with $\lambda_i > 0$ for $i = 1, \ldots, d$. It follows that

$$Jac(T_{p,t})(x) = \det((1-t)Id + t\tilde{\nabla} T_p(x)) = \prod_{j=1}^{d} \big((1-t) + \lambda_j t\big).$$

Passing to logarithms, we have by concavity

$$\log(Jac(T_{p,t})(x)) \geq t\log(Jac(T_p)(x)) + (1-t)\log(Jac(Id)) = t\log(Jac(T_p)(x)).$$

Hence

$$Jac(T_{p,t})(x) \geq \Big(Jac(T_p)\Big)^t(x) \geq \left(\frac{\rho_1(x)}{\rho_2(T_p(x))}\right)^t \exp\Big\{ \sum_{j=1}^{d} -\frac{1}{2}t(x_j^2 - T_p(x)_j^2) \Big\}. \tag{2.5.1}$$

Step 3. We have the following

$$
\begin{aligned}
&\exp\Big\{\sum_{j=1}^{d} -\frac{1}{2}(T_{p,t}(x)_j^2 - x_j^2)\Big\} Jac(T_{p,t})(x)\\
&\qquad \geq \exp\Big\{\sum_{j=1}^{d} -\frac{1}{2}(T_{p,t}(x)_j^2 - x_j^2)\Big\} \Big(\frac{\rho_1(x)}{\rho_2(T_p(x))}\Big)^t \exp\Big\{\sum_{j=1}^{d} -\frac{1}{2}t(x_j^2 - T_p(x)_j^2)\Big\}\\
&\qquad = \Big(\frac{\rho_1(x)}{\rho_2(T_p(x))}\Big)^t \exp\Big\{\sum_{j=1}^{d} -\frac{1}{2}(T_{p,t}(x)_j^2 - x_j^2 + t x_j^2 - tT_p(x)_j^2)\Big\}\\
&\qquad = \Big(\frac{\rho_1(x)}{\rho_2(T_p(x))}\Big)^t \exp\Big\{\sum_{j=1}^{d} -\frac{1}{2}\Big(((1-t)x_j + tT_p(x)_j)^2 - ((1-t)x_j^2 + tT_p(x)_j^2\Big)\Big\}\\
&\qquad = \Big(\frac{\rho_1(x)}{\rho_2(T_p(x))}\Big)^t \exp\Big\{\sum_{j=1}^{d} -\frac{1}{2}\big(x_j - T_p(x)_j\big)^2(t^2 - t)\Big\}\\
&\qquad = \Big(\frac{\rho_1(x)}{\rho_2(T_p(x))}\Big)^t \exp\Big\{ -\frac{1}{2}\|x - T_p(x)\|_d^2(t^2 - t)\Big\}.
\end{aligned}
$$

Thus

$$
\begin{aligned}
\gamma(T_{p,t}(A)) &= \int_A Jac(T_{p,t})(x) \prod_{j=1}^{d} \frac{1}{\sqrt{2\pi}} \exp\Big\{-\frac{1}{2}T_{p,t}(x)_j^2\Big\} \mathcal{L}^d(dx)\\
&= \int_A Jac(T_{p,t})(x) \exp\Big\{\sum_{j=1}^{d} -\frac{1}{2}(T_{p,t}(x)_j^2 - x_j^2)\Big\}\gamma(dx)\\
&\geq \int_A \Big(\frac{\rho_1(x)}{\rho_2(T_p(x))}\Big)^t \exp\Big\{\frac{1}{2}\|x - T_p(x)\|_d^2(t - t^2)\Big\}\gamma(dx)\\
&\geq \frac{1}{C^t}\int_A \rho_1(x)^t \gamma(dx)\\
&\geq \frac{1}{C^t}\int_A \rho_1(x)^{t-1} \mu(dx)\\
&\geq \frac{1}{C}\mu(A).
\end{aligned}
$$

Step 4. Since $(Id, T_p)_\sharp \mu_d \rightharpoonup (Id, T)_\sharp \mu_d$ as $p \searrow 1$, see Theorem 7.1 of [5], using the techniques of the proof of Theorem 2.17, one can prove that

$$\gamma_d(T_t(\Gamma_d \cap A \times X)) \geq \frac{1}{C}\mu_d(A).$$

□

Remark 2.19. We summarize the results obtained so far. If $\rho_1, \rho_2 \leq C$, then from (2.4.2) it follows that the densities of μ_d and ν_d enjoy the same property with the same constant C. Identifying μ_d, ν_d and γ_d with the corresponding measures on $\mathbb{R}^d$, we have from Theorem 2.18:

$$\gamma_d\big(T_t(\Gamma_d \cap A \times X)\big) \geq \frac{1}{C}\mu_d(A), \qquad \forall A \in \mathcal{B}(X), t \in [0,1].$$

From Theorem 2.17 we have the same inequality for the ∞-dimensional measures:

$$\gamma\big(T_t(\Gamma \cap A \times X)\big) \geq \frac{1}{C}\mu(A), \qquad \forall A \in \mathcal{B}(X)_\gamma, t \in [0,1].$$

As Proposition 2.14 and Proposition 2.15 show, this estimate implies $\mu(a(\mathcal{T}))=0$ and the continuity of the conditional probabilities μ_y. Since the optimal finite dimensional map T_d is invertible, following the argument of Theorem 2.18 we can also prove

$$\gamma\big(T_{1-t}(\Gamma\cap X\times A)\big)\geq \frac{1}{C}\nu(A), \tag{2.5.2}$$

and adapting the proofs of Proposition 3.16 and Proposition 3.17 we can prove that $\nu(b(\mathcal{T}))=0$ and the continuity of the conditional probabilities ν_y. So we have

$$\mu=\int_{\mathcal{S}}\mu_y m(dy),\quad \nu=\int_{\mathcal{S}}\nu_y m(dy),\qquad \mu_y,\nu_y \text{ continuous for } m-a.e.y\in\mathcal{S}.$$

In the next Section we remove the hypothesis $\rho_1,\rho_2\leq M$.

2.6 Solution

Let π be the weak limit of $\pi_d=(Id,T_d)_\sharp\mu_d$ and Γ any $\|\cdot\|_{H(\gamma)}$-cyclically monotone set such that $\pi(\Gamma)=1$. All the definition of Section 3.1 are referred to this Γ.

Proposition 2.20. *Let $\mu,\nu\in\mathcal{P}(X)$ be such that $\mu,\nu\ll\gamma$. Then $\mu(a(\mathcal{T}))=\nu(b(\mathcal{T}))=0$.*

Proof. Let $\mu=\rho_1\gamma$ and $\nu=\rho_2\gamma$. We only prove that $\mu(a(\mathcal{T}))=0$. The other statement follows similarly.

Step 1. Assume by contradiction that $\mu(a(\mathcal{T}))>0$. Let $A\subset a(\mathcal{T})$ be such that $\mu(A)>0$ and for every $x\in A$, $\rho_1(x)\leq M$ for some positive constant M. Consider $\gamma_{\llcorner\mathcal{T}}$ and its disintegration

$$\gamma_{\llcorner\mathcal{T}}=\int_{\mathcal{S}}\gamma_y m_\gamma(dy),\quad \gamma_y(\mathcal{T})=1,\ m_\gamma-a.e.y\in\mathcal{S}.$$

Consider the initial point map $a:\mathcal{S}\to A$ and the measure $a_\sharp m_\gamma$. Observe that since

$$\forall B\subset A:\mu(B)>0\quad\Rightarrow\quad \gamma(R(B)\cap\mathcal{T})>0,$$

it follows that $\mu_{\llcorner A}\ll a_\sharp m_\gamma$. Hence there exists $\hat{A}\subset A$ of positive $a_\sharp m_\gamma$-measure such that the map

$$\hat{A}\ni x\mapsto h(x):=\frac{d\mu_{\llcorner A}}{da_\sharp m_\gamma}(x)$$

verifies $h(x)\leq M'$ for some positive constan M'.

Step 2. Considering

$$\mu_{\llcorner\hat{A}},\qquad \hat{\gamma}:=\int_{R(\hat{A})\cap\mathcal{S}}h(a(y))\gamma_y m_\gamma(dy),$$

we have the claim. Indeed both have uniformly bounded densities w.r.t. γ and $\mathcal{T}_e$ is still a transport set for the transport problem between $\mu_{\llcorner\hat{A}}$ and $\hat{\gamma}$. Indeed for $S\subset\mathcal{S}$

$$\begin{aligned}\mu_{\llcorner\hat{A}}(\cup_{y\in S}R(y))&=\mu_{\llcorner\hat{A}}(a(S))\\&=\int_{a(S)}h(a)(a_\sharp m_\gamma)(da)\\&=\int_S h(a(y))m_\gamma(dy)=\hat{\gamma}(\cup_{y\in S}R(y)).\end{aligned}$$

Hence we can project the measures, obtain the finite dimensional estimate of Theorem 2.18, obtain the infinite dimensional estimate through Theorem 2.17 and finally by Proposition 2.14 get that $\mu(\hat{A})=0$, that is a contradiction with $\mu(\hat{A})>0$. In the same way, following Remark 2.19, we obtain that $\nu(b(\mathcal{T}))=0$. □

It follows that the disintegration formula (2.2.5) holds true on the whole transportation set:

$$\mu = \int \mu_y m(dy), \qquad \nu = \int \nu_y m(dy).$$

Proposition 2.21. *For m-a.e. $y \in \mathcal{S}$ the conditional probabilities μ_y and ν_y have no atoms.*

Proof. We only prove the claim for μ_y.

Step 1. Suppose by contradiction that there exist a measurable set $\hat{\mathcal{S}} \subset \mathcal{S}$ such that $m(\hat{\mathcal{S}}) > 0$ and for every $y \in \hat{\mathcal{S}}$ there exists $x(y)$ such that $\mu_y(\{x(y)\}) > 0$. Restrict and normalize both μ and ν to $R(\hat{\mathcal{S}})$, and denote them again with μ and ν.

Consider the sets $K_{i,M} := \{x \in X : \rho_i \leq M\}$ for $i = 1, 2$. Note that $\mu(K_{1,M}) \geq 1 - c_1(M)$ and $\nu(K_{2,\delta}) \geq 1 - c_2(M)$ with $c_i(M) \to 0$ as $M \nearrow +\infty$. Hence for M sufficiently large the conditional probabilities of the disintegration of $\mu \llcorner_{K_{1,M}}$ have atoms, therefore we can assume, possibly restricting $\hat{\mathcal{S}}$, that for all $y \in \hat{\mathcal{S}}$ it holds $x(y) \in K_{1,M}$.

Step 2. Define

$$\mu_{y,M} := \mu_y \llcorner_{K_{1,M}}, \qquad \nu_{y,M} := \nu_y \llcorner_{K_{2,M}},$$

and introduce the set

$$D(N) := \Big\{ y \in \hat{\mathcal{S}} : \frac{\mu_{y,M}(R(y))}{\nu_{y,M}(R(y))} \leq N \Big\}.$$

Then for sufficiently large N, $m(D(N)) > 0$. The map $D(N) \ni y \mapsto h(y) := \nu_{y,M}(R(y))/\mu_{y,M}(R(y)) \leq N$ permits to define

$$\hat{\mu} := \int_{D(N)} h(y) \mu_{y,M} m(dy), \qquad \hat{\nu} := \nu \llcorner_{R(D(N)) \cap K_{2,M}}.$$

It follows that $\hat{\mu}$ and $\hat{\nu}$ have bounded densities w.r.t. γ and the set $\hat{\mathcal{T}} := \mathcal{T} \cap G(K_{1,\delta}) \cap G^{-1}(K_{2,\delta})$ is a transport set for the transport problem between $\hat{\mu}$ and $\hat{\nu}$.

It follows from Theorem 2.17 and Theorem 2.18 that $\hat{\gamma} := \gamma \llcorner_{\hat{\mathcal{T}}}$ verifies Assumption 5 w.r.t. $G \cap K_{1,M} \times X \cap X \times K_{2,M}$. Therefore from Proposition 3.17 follows that the conditional probabilities $\hat{\mu}_y$ of the disintegration of $\hat{\mu}$ are continuous. Since $\hat{\mu}_y = c(y) \mu_y \llcorner_{\hat{\mathcal{T}}}$ for some positive constant $c(y)$, we have a contradiction. □

It follows straightforwardly the existence of an optimal invertible transport map.

Theorem 2.22. *Let $\mu, \nu \in \mathcal{P}(X)$ absolute continuous w.r.t. γ and assume that there exists $\pi \in \Pi(\mu, \nu)$ such that $\mathcal{I}(\pi)$ is finite. Then there exists a solution for the Monge minimization problem*

$$\min_{T : T_\sharp \mu = \nu} \int_X \|x - T(x)\|_{H(\gamma)} \mu(dx).$$

Moreover we can find T μ-essentially invertible.

Proof. For m-a.e $y \in \mathcal{S}$ μ_y and ν_y are continuous. Since $R(y)$ is one dimensional and the ray map $\mathbb{R} \ni t \mapsto g(t, y)$ is an isometry w.r.t. $\|\cdot\|_{H(\gamma)}$, we can define the non atomic measures $g(y, \cdot)_\sharp \mu_y, g(y, \cdot)_\sharp \nu_y \in \mathcal{P}(\mathbb{R})$. By the one-dimensional theory, there exists a monotone map $T_y : \mathbb{R} \to \mathbb{R}$ such that

$$T_{y\,\sharp}\Big(g(y, \cdot)_\sharp \mu_y\Big) = g(y, \cdot)_\sharp \nu_y.$$

Using the inverse of the ray map, we can define T_y on $R(y)$. Hence for m-a.e. $y \in \mathcal{S}$ we have a $\|\cdot\|_{H(\gamma)}$-cyclically monotone map T_y such that $T_{y\,\sharp} \mu_y = \nu_y$. To conclude define $T : \mathcal{T} \to \mathcal{T}$ such that $T = T_y$ on $R(y)$. Indeed T is μ-measurable, μ-essentially invertible and $T_\sharp \mu = \nu$. For the details, see the proof of Theorem 6.2 of [9]. □

Chapter 3

The Monge problem for general geodesic distance cost

In this chapter we study the following Monge minimization problem: given two Borel probability measure $\mu, \nu \in \mathcal{P}(X)$, where (X, d) is a Polish space, we study the minimization of the functional

$$\mathcal{I}(T) = \int d_N(x, T(x))\mu(dy)$$

where T varies over all Borel transport maps and d_N is a Borel distance taking also the value ∞ that makes (X, d_N) a possibly branching geodesic space. We will apply the results to the obstacle problem: let $C \subset \mathbb{R}^d$ be a convex set with $\partial C = M$ smooth, $(d-1)$-dimensional compact submanifold of $\mathbb{R}^d$. Let $X = (\mathbb{R}^d \setminus C) \cup M$, $\mu, \nu \in \mathcal{P}(X)$ and $d_M(x, y)$ be the infimum among all the Lipschitz curves in X connecting to x to y of the euclidean length of such curves. We will prove the existence of a solution for

$$\min_{T: T_\sharp \mu = \nu} \int d_M(x, T(x))\mu(dx),$$

provided $\mu \ll \mathcal{L}^d$.

Chapter 3 is organized as follows.

Section 3.1 shows how using only the d_N-cyclical monotonicity of a set Γ we can obtain a partial order relation $G \subset X \times X$ as follows (Lemma 3.3 and Proposition 3.9): xGy iff there exists $(w, z) \in \Gamma$ and a geodesic γ, passing trough w and z and with direction $w \to z$, such that x, y belongs to γ and $\gamma^{-1}(x) \leq \gamma^{-1}(y)$. This set G is analytic, and allows to define

- the transport rays set R (Definition 2.6),
- the transport sets $\mathcal{T}_e$, $\mathcal{T}$ (with and without and points) (3.1.4),
- the set of initial points a and final points b (3.1.7).

Even if this part of Section 3.1 contains the same results of the first part of Section 1.2 of Chapter 1, we show their proofs again to underline what depends and what doesn't depend on the branching property of the space. The main difference with the non-branching case is that here R is not an equivalence relation. Therefore the approach proposed in Chapter 1 doesn't work anymore and indeed the only common part with Chapter 1 is the first part of Section 3.1.

To obtain an equivalence relation $H \subset X \times X$ we have to consider the set of couples (x, y) for $x, y \in \mathcal{T}$ such that there is a continuous path from x to y, union of a finite number of transport rays never passing through $a \cup b$, Definition 3.8. In Proposition 3.9 we prove that H is an equivalence relation.

This chapter is based on the work [15].

Section 3.2 proves that the compatibility conditions (1.a) and (1.b) between d_N and d imply that the disintegration induced by H on $\mathcal{T}$ is strongly consistent (Proposition 3.14). Using this fact we can reduce the analysis on $H(y)$ for y in the quotient set.

In Section 3.3 we prove Theorem 0.6. We first introduce the operation $A \mapsto A_t$, the translation along geodesics (3.3.1), and show that $t \mapsto \mu(A_t)$ is a $\mathcal{A}$-measurable function if A is analytic (Lemma 3.15). Next, we show that under the assumption

$$\mu(A) > 0 \quad \Longrightarrow \quad \mu(A_{t_n}) \geq C\mu(A)$$

for an infinitesimal sequence t_n and $C > 0$, the set of initial points a is μ-negligible (Proposition 3.16) and the conditional probabilities μ_y are continuous.

In Section 3.4 we prove Theorem 0.7. First in Theorem 3.18 we prove that gluing all the d_N-cyclically monotone maps defined on $H(y)$ we obtain a measurable transference map T from μ to ν d_N-cyclically monotone. Then the assumption on the structure of Γ is stated (Assumption 8) and in Proposition 3.19 we show that on the equivalence class $H(y)$ satisfying Assumption 8 there exists an optimal transference map T_y from μ_y to ν_y, provided the quotient measure and the marginal probabilities of μ_y induced by the partition given by Assumption 8 are continuous.

Section 3.5 gives an application of Theorem 0.7: M is a connected smooth hyper-surface of $\mathbb{R}^d$ that is the boundary of a convex and compact set C. Let $X = cl(\mathbb{R}^d \setminus C)$. The distance d_M is the minimum of the euclidean length among all the Lipschitz curves in X (3.5.1). Hence C is to be intended as an obstacle for euclidean geodesics. The geodesic space (X, d_M) fits into the setting of Theorem 0.7 (Lemma 3.22 and Remark 3.24).

If $\mu \ll \mathcal{L}^d$ then the μ-measure of the set of initial points is zero and the marginal μ_y are continuous (Lemma 3.25). Finally we show in Proposition 3.27 and Proposition 3.29 that any d_M-cyclically monotone set and μ satisfy the hypothesis of Proposition 3.19. It follows the existence of a solution for the Monge minimization problem.

3.1 Optimal transportation in geodesic spaces

From now on we assume the following:

1. (X, d) Polish space;
2. $d_N : X \times X \to [0, +\infty]$ is a Borel distance;
3. (X, d_N) is a geodesic space;

Since we have two metric structures on X, we denote the quantities relating to d_N with the subscript N: for example

$$B_r(x) = \{y : d(x, y) < r\}, \quad B_{r,N}(x) = \{y : d_N(x, y) < r\}.$$

In particular we will use the notation

$$D_N(x) = \{y : d_N(x, y) < +\infty\},$$

$(\mathcal{K}, d_H)$ for the compact sets of (X, d) with the Hausdorff distance d_H and $(\mathcal{K}_N, d_{H,N})$ for the compact sets of (X, d_N) with the Hausdorff distance $d_{H,N}$. We recall that $(\mathcal{K}, d_H)$ is Polish.

Let $\mu, \nu \in \mathcal{P}(X)$ and consider the transportation problem with cost $c(x, y) = d_N(x, y)$, and let $\pi \in \Pi(\mu, \nu)$ be a d_N-cyclically monotone transference plan with finite cost. By inner regularity, we can assume that the optimal transference plan is concentrated on a σ-compact d_N-cyclically monotone set $\Gamma \subset \{d_N(x, y) < +\infty\}$.

Consider the set

$$\Gamma' := \Big\{ (x,y) : \exists I \in \mathbb{N}_0, (w_i, z_i) \in \Gamma \text{ for } i = 0, \dots, I,\ z_I = y$$
$$w_{I+1} = w_0 = x,\ \sum_{i=0}^{I} d_N(w_{i+1}, z_i) - d_N(w_i, z_i) = 0 \Big\}. \tag{3.1.1}$$

In other words, we concatenate points $(x,z), (w,y) \in \Gamma$ if they are initial and final point of a cycle with total cost 0.

Lemma 3.1. *The following holds:*

1. $\Gamma \subset \Gamma' \subset \{d_N(x,y) < +\infty\}$;
2. *if* Γ *is analytic, so is* Γ';
3. *if* Γ *is* d_N*-cyclically monotone, so is* Γ'.

Proof. For the first point, set $I = 0$ and $(w_{n,0}, z_{n,0}) = (x,y)$ for the first inclusion. If $d_N(x,y) = +\infty$, then $(x,y) \notin \Gamma$ and all finite set of points in Γ are bounded.

For the second point, observe that

$$\begin{aligned}\Gamma' &= \bigcup_{I \in \mathbb{N}_0} P_{1,2I+1}(A_I) \\ &= \bigcup_{I \in \mathbb{N}_0} P_{1,2I+1}\bigg(\prod_{i=0}^{I} \Gamma \cap \bigg\{ \prod_{i=0}^{I} (w_i, z_i) : \sum_{i=0}^{I} d_N(w_{i+1}, z_i) - d_N(w_i, z_i) = 0, w_{I+1} = w_0 \bigg\}\bigg).\end{aligned}$$

For each $I \in \mathbb{N}_0$, since d_N is Borel, it follows that

$$\bigg\{ \prod_{i=0}^{I} (w_i, z_i) : \sum_{i=0}^{I} d_N(w_{i+1}, z_i) - d_N(w_i, z_i) = 0, w_{I+1} = w_0 \bigg\}$$

is Borel in $\prod_{i=0}^{I}(X \times X)$, so that for Γ analytic each set $A_{n,I}$ is analytic. Hence $P_{1,2I+1}(A_I)$ is analytic, and since the class Σ^1_1 is closed under countable unions and intersections it follows that Γ' is analytic.

For the third point, observe that for all $(x_j, y_j) \in \Gamma'$, $j = 0, \dots, J$, there are $(w_{j,i}, z_{j,i}) \in \Gamma$, $i = 0, \dots, I_j$, such that

$$d_N(x_j, y_j) + \sum_{i=0}^{I_j - 1} d_N(w_{j,i+1}, z_{j,i}) - \sum_{i=0}^{I_j} d_N(w_{j,i}, z_{j,i}) = 0.$$

Hence we can write for $x_{J+1} = x_0$, $w_{j,I_j+1} = w_{j+1,0}$, $w_{J+1,0} = w_{0,0}$

$$\sum_{j=0}^{J} d_N(x_{j+1}, y_j) - d_N(x_j, y_j) = \sum_{j=0}^{J} \sum_{i=0}^{I_j} d_N(w_{j,i+1}, z_{j,i}) - d_N(w_{j,i}, z_{j,i}) \geq 0,$$

using the d_N-cyclical monotonicity of Γ. □

Definition 3.2 (Transport rays)**.** Define the *set of oriented transport rays*

$$G := \Big\{ (x,y) : \exists (w,z) \in \Gamma', d_N(w,x) + d_N(x,y) + d_N(y,z) = d_N(w,z) \Big\}. \tag{3.1.2}$$

For $x \in X$, the *outgoing transport rays from* x is the set $G(x)$ and the *incoming transport rays in* x is the set $G^{-1}(x)$. Define the *set of transport rays* as the set

$$R := G \cup G^{-1}. \tag{3.1.3}$$

The set G is the set of all couples of points on oriented geodesics with endpoints in Γ'. In R the couples are non oriented.

Lemma 3.3. *The following holds:*

1. *G is d_N-cyclically monotone;*
2. *$\Gamma' \subset G \subset \{d_N(x,y) < +\infty\}$;*
3. *the sets G, $R := G \cup G^{-1}$ are analytic.*

Proof. The second point follows by the definition: if $(x,y) \in \Gamma'$, just take $(w,z) = (x,y)$ in the r.h.s. of (3.1.2).

The third point is consequence of the fact that

$$G = P_{34}\Big(\big(\Gamma' \times X \times X\big) \cap \Big\{(w,z,x,y) : d_N(w,x) + d_N(x,y) + d_N(y,z) = d_N(w,z)\Big\}\Big),$$

and the result follows from the properties of analytic sets.

The first point follows from the following observation: if $(x_i, y_i) \in \gamma_{[w_i,z_i]}$, then from triangle inequality

$$\begin{aligned} d_N(x_{i+1}, y_i) - d_N(x_i, y_i) &\geq d_N(x_{i+1}, z_i) - d_N(z_i, y_i) - d_N(x_i, y_i) \\ &= d_N(x_{i+1}, z_i) - d_N(x_i, z_i) \\ &\geq d_N(w_{i+1}, z_i) - d_N(w_{i+1}, x_{i+1}) - d_N(x_i, z_i) \\ &= d_N(w_{i+1}, z_i) - d_N(w_i, z_i) + d_N(w_i, x_i) - d_N(w_{i+1}, x_{i+1}). \end{aligned}$$

Since $(w_{n+1}, x_{n+1}) = (w_1, x_1)$, it follows that

$$\sum_{i=1}^{n} d_N(x_{i+1}, y_i) - d_N(x_i, y_i) \geq \sum_{i=1}^{n} d_N(w_{i+1}, z_i) - d_N(w_i, z_i) \geq 0.$$

Hence the set G is d_N-cyclically monotone. □

Definition 3.4. Define the *transport sets*

$$\mathcal{T} := P_1\big(G^{-1} \setminus \{x = y\}\big) \cap P_1\big(G \setminus \{x = y\}\big), \tag{3.1.4a}$$

$$\mathcal{T}_e := P_1\big(G^{-1} \setminus \{x = y\}\big) \cup P_1\big(G \setminus \{x = y\}\big). \tag{3.1.4b}$$

Since G and G^{-1} are analytic sets, $\mathcal{T}$, $\mathcal{T}_e$ are analytic. The subscript e refers to the endpoints of the geodesics: clearly we have

$$\mathcal{T}_e = P_1(R \setminus \{x = y\}). \tag{3.1.5}$$

The following lemma shows that we have only to study the Monge problem in $\mathcal{T}_e$.

Lemma 3.5. *It holds $\pi(\mathcal{T}_e \times \mathcal{T}_e \cup \{x = y\}) = 1$.*

Proof. If $x \in P_1(\Gamma \setminus \{x = y\})$, then $x \in G^{-1}(y) \setminus \{y\}$. Similarly, $y \in P_2(\Gamma \setminus \{x = y\})$ implies that $y \in G(x) \setminus \{x\}$. Hence $\Gamma \setminus \mathcal{T}_e \times \mathcal{T}_e \subset \{x = y\}$. □

As a consequence, $\mu(\mathcal{T}_e) = \nu(\mathcal{T}_e)$ and any maps T such that for $\nu \llcorner_{\mathcal{T}_e} = T_\sharp \mu \llcorner_{\mathcal{T}_e}$ can be extended to a map T' such that $\nu = T_\sharp \mu$ with the same cost by setting

$$T'(x) = \begin{cases} T(x) & x \in \mathcal{T}_e \\ x & x \notin \mathcal{T}_e \end{cases} \tag{3.1.6}$$

Definition 3.6. Define the multivalued *endpoint graphs* by:

$$a := \big\{(x,y) \in G^{-1} : G^{-1}(y) \setminus \{y\} = \emptyset\big\}, \tag{3.1.7a}$$
$$b := \big\{(x,y) \in G : G(y) \setminus \{y\} = \emptyset\big\}. \tag{3.1.7b}$$

We call $P_2(a)$ the set of *initial points* and $P_2(b)$ the set of *final points*.

Proposition 3.7. *The following holds:*

1. *the sets*
$$a, b \subset X \times X, \quad a(A), b(A) \subset X,$$
belong to the $\mathcal{A}$-class if A analytic;
2. $a \cap b \cap \mathcal{T}_e \times X = \emptyset$*;*
3. $a(\mathcal{T}) = a(\mathcal{T}_e)$*,* $b(\mathcal{T}) = b(\mathcal{T}_e)$*;*
4. $\mathcal{T}_e = \mathcal{T} \cup a(\mathcal{T}) \cup b(\mathcal{T})$*,* $\mathcal{T} \cap (a(\mathcal{T}) \cup b(\mathcal{T})) = \emptyset$*.*

Proof. Define

$$C := \big\{(x,y,z) \in \mathcal{T}_e \times \mathcal{T}_e \times \mathcal{T}_e : y \in G(x), z \in G(y)\big\} = (G \times X) \cap (X \times G) \cap \mathcal{T}_e \times \mathcal{T}_e \times \mathcal{T}_e,$$

that is clearly analytic. Then

$$b = \big\{(x,y) \in G : y \in G(x), G(y) \setminus \{y\} = \emptyset\big\} = G \setminus P_{12}(C \setminus X \times \{y = z\}),$$

$$b(A) = \big\{y : y \in G(x), G(y) \setminus \{y\} = \emptyset, x \in A\big\} = P_2(G \cap A \times X) \setminus P_2(C \setminus X \times \{y = z\}).$$

A similar computation holds for a:

$$a = G^{-1} \setminus P_{23}(C \setminus \{x = y\} \times X), \quad a(A) = P_1(G_S \cap X \times A) \setminus P_1(C \setminus \{x = y\} \times X).$$

Hence $a, b \in \mathcal{A}(X \times X)$, $a(A), b(A) \in \mathcal{A}(X)$, being the intersection of an analytic set with a coanalytic one. If $x \in \mathcal{T}_e \setminus \mathcal{T}$, then it follows that $G(x) = \{x\}$ or $G^{-1}(x) = \{x\}$ hence $x \in a(x) \cup b(x)$.

The other points follow easily. □

Definition 3.8 (Chain of transport rays)**.** Define the set of *chain of transport rays*

$$\begin{aligned} H := \bigg\{ &(x,y) \in \mathcal{T}_e \times \mathcal{T}_e : \exists I \in \mathbb{N}_0, z_i \in \mathcal{T} \text{ for } 1 \leq i \leq I, \\ &(z_i, z_{i+1}) \in R,\ 0 \leq i \leq I+1,\ z_0 = x,\ z_{I+1} = y \bigg\}. \end{aligned} \tag{3.1.8}$$

Using similar techniques of Lemma 3.1 it can be shown that H is analytic.

Proposition 3.9. *The set $H \cap \mathcal{T} \times \mathcal{T}$ is an equivalence relation on $\mathcal{T}$. The set G is a partial order relation on $\mathcal{T}_e$.*

Proof. Using the definition of H, one has in $\mathcal{T}$:

1. $x \in \mathcal{T}$ clearly implies that $(x,x) \in H$;
2. since R is symmetric, if $y \in H(x)$ then $x \in H(y)$;
3. if $y \in H(x)$, $z \in H(y)$, $x, y, z \in \mathcal{T}$. Glue the path from x to y to the one from y to z. Since $y \in \mathcal{T}$, $z \in H(x)$.

The second part follows similarly:

1. $x \in \mathcal{T}_e$ implies that
$$\exists (x,y) \in \big(G \setminus \{x=y\}\big) \cup \big(G^{-1} \setminus \{x=y\}\big),$$
so that in both cases $(x,x) \in G$;

2. $(x,y),(y,z) \in G \setminus \{x=y\}$ implies by d_N-cyclical monotonicity that $(x,z) \in G$.

□

We finally show that we can assume that the μ-measure of final points and the ν-measure of the initial points are 0.

Lemma 3.10. *The sets $G \cap b(\mathcal{T}) \times X$, $G \cap X \times a(\mathcal{T})$ is a subset of the graph of the identity map.*

Proof. From the definition of b one has that
$$x \in b(\mathcal{T}) \quad \Longrightarrow \quad G(x) \setminus \{x\} = \emptyset,$$
A similar computation holds for a. □

Hence we conclude that
$$\pi(b(\mathcal{T}) \times X) = \pi(G \cap b(\mathcal{T}) \times X) = \pi(\{x=y\})$$
and following (3.1.6) we can assume that
$$\mu(b(\mathcal{T})) = \nu(a(\mathcal{T})) = 0.$$

3.2 Partition of the transport set

To perform a disintegration we have to assume some regularity of the support Γ of the transport plan $\pi \in \Pi(\mu,\nu)$. From now on we will assume the following:

Assumption 6. We say that Γ satisfies Assumption 6 if

(a) for all $x \in \mathcal{T}$ and for all $r > 0$ the set $H(x) \cap \overline{B_{r,N}(x)}^{d_N}$ is d-closed;

(b) for all $x \in \mathcal{T}$ there exists $r > 0$ such that $d_N(x,\cdot)_{\llcorner H(x) \cap \overline{B_r(x)}}$ is bounded.

Note that points (a) and (b) of Assumption 6 were already introduced at page 12. Let $\{x_i\}_{i \in \mathbb{N}}$ be a dense sequence in (X,d).

Lemma 3.11. *The sets*
$$W_{ijk} := \Big\{ x \in \mathcal{T} \cap \bar{B}_{2^{-j}}(x_i) : d_N(x,\cdot)_{\llcorner H(x) \cap \bar{B}_{2^{-j}}(x_i)} \leq k \Big\}$$
form a countable covering of $\mathcal{T}$ of class $\mathcal{A}$.

Proof. We first prove the measurability. We consider separately the conditions defining W_{ijk}.

Point 1. The set
$$A_{ij} := \mathcal{T} \cap \bar{B}_{2^{-j}}(x_i)$$
is clearly analytic.

Point 2. The set
$$D_{ijk} := \Big\{ (x,y) \in H : d(x_i,y) \leq 2^{-j}, d_N(x,y) > k \Big\}$$

is again analytic. We finally can write

$$W_{ijk} = A_{ij} \cap P_1(D_{ijk})^c,$$

and the fact that $\mathcal{A}$ is a σ-algebra proves that $W_{ijk} \in \mathcal{A}$.

To show that it is a covering, notice that from (b) of Assumption 6 for all $x \in \mathcal{T}$ there exists $r > 0$ such that, on the set $H(x) \cap \bar{B}_r(x)$, $d_N(x,\cdot)$ is bounded. Choose j and i such that $2^{-j-1} \leq r$ and $d(x_i, x) \leq 2^{-j-1}$, hence

$$\bar{B}_{2^{-j}}(x_i) \subset \bar{B}_r(x)$$

and therefore for some $\bar{k} \in \mathbb{N}$ we obtain that $x \in W_{ijk}$. □

Remark 3.12. Observe that $\bar{B}_{2^{-j}}(x_i) \cap H(x)$ is closed for all $x \in W_{ijk}$.

Indeed take $\{y_n\}_{n\in\mathbb{N}} \subset \bar{B}_{2^{-j}}(x_i) \cap H(x)$ with $d(y_n, y) \to 0$ as $n \to +\infty$, then since $x \in W_{ijk}$ it holds $d_N(x, y_n) \leq k$. By (a) of Assumption 6, $d_N(x, y) \leq k$ and $y \in \bar{B}_{2^{-j}}(x_i) \cap H(x)$.

Lemma 3.13. *There exist μ-negligible sets $N_{ijk} \subset W_{ijk}$ such that the family of sets*

$$\mathcal{T}_{ijk} = H^{-1}(W_{ijk} \setminus N_{ijk}) \cap \mathcal{T}$$

is a countable covering of $\mathcal{T} \setminus \cup_{ijk} N_{ijk}$ into saturated analytic sets.

Proof. First of all, since $W_{ijk} \in \mathcal{A}$, then there exists μ-negligible set $N_{ijk} \subset W_{ijk}$ such that $W_{ijk} \setminus N_{ijk} \in \mathcal{B}(X)$. Hence $\{W_{ijk} \setminus N_{ijk}\}_{i,j,k\in\mathbb{N}}$ is a countable covering of $\mathcal{T} \setminus \cup_{ijk} N_{ijk}$. It follows immediately that $\{\mathcal{T}_{ijk}\}_{i,j,k\in\mathbb{N}}$ satisfies the lemma. □

From any analytic countable covering, we can find a countable partition into $\mathcal{A}$-class saturated sets by defining

$$\mathcal{Z}_m := \mathcal{T}_{i_m j_m k_m} \setminus \bigcup_{m'=1}^{m-1} \mathcal{T}_{i_{m'} j_{m'} k_{m'}}, \tag{3.2.1}$$

where

$$\mathbb{N} \ni m \mapsto (i_m, j_m, k_m) \in \mathbb{N}^3$$

is a bijective map. Since H is an equivalence relation on $\mathcal{T}$, we use this partition to prove the strong consistency.

On $\mathcal{Z}_m$, $m > 0$, we define the closed valued map

$$\mathcal{Z}_m \ni x \mapsto F(x) := H(x) \cap \bar{B}_{2^{-j_m}}(x_{i_m}). \tag{3.2.2}$$

Proposition 3.14. *There exists a μ-measurable cross section $f : \mathcal{T} \to \mathcal{T}$ for the equivalence relation H.*

Proof. First we show that F is $\mathcal{A}$-measurable: for $\delta > 0$,

$$\begin{aligned} F^{-1}(B_\delta(y)) &= \Big\{x \in \mathcal{Z}_m : H(x) \cap B_\delta(y) \cap \bar{B}_{2^{-j_m}}(x_{i_m}) \neq \emptyset\Big\} \\ &= \mathcal{Z}_m \cap P_1\Big(H \cap \big(X \times B_\delta(y) \cap \bar{B}_{2^{-j_m}}(x_{i_m})\big)\Big). \end{aligned}$$

Being the intersection of two $\mathcal{A}$-class sets, $F^{-1}(B_\delta(y))$ is in $\mathcal{A}$. In Remark 3.12 we have observed that F is a closed-valued map, hence, from Lemma 5.1.4 of [25], graph(F) is $\mathcal{A}$-measurable.

By Corollary 5.7 there exists a $\mathcal{A}$-class section $f_m : \mathcal{Z}_m \to \bar{B}_{2^{-j_m}}(x_{i_m})$. The proposition follows by setting $f\llcorner_{\mathcal{Z}_m} = f_m$ on $\cup_m \mathcal{Z}_m$, and defining it arbitrarily on $\mathcal{T} \setminus \cup_m \mathcal{Z}_m$: the latter being μ-negligible, f is μ-measurable. □

Up to a μ-negligible saturated set $\mathcal{T}_N$, we can assume it to have σ-compact range: just let $S \subset f(\mathcal{T})$ be a σ-compact set where $f_\sharp \mu$ is concentrated, and set

$$\mathcal{T}_S := H^{-1}(S) \cap \mathcal{T}, \quad \mathcal{T}_N := \mathcal{T} \setminus \mathcal{T}_S, \quad \mu(\mathcal{T}_N) = 0. \tag{3.2.3}$$

Hence we have a measurable cross-section

$$\mathcal{S} := S \cup f(\mathcal{T}_N) = (\text{Borel}) \cup (f(\mu\text{-negligible})).$$

Hence Disintegration Theorem 5.3 yields

$$\mu \llcorner_{\mathcal{T}} = \int_S \mu_y m(dy), \quad m = f_\sharp \mu \llcorner_{\mathcal{T}}, \ \mu_y \in \mathcal{P}(H(y)) \tag{3.2.4}$$

and the disintegration is strongly consistent since the quotient map $f : \mathcal{T} \to \mathcal{T}$ is μ-measurable and $(\mathcal{T}, \mathcal{B}(\mathcal{T}))$ is countably generated.

Observe that H induces an equivalence relation also on $\mathcal{T} \times X \cap \Gamma$ where the equivalence classes are $H(y) \cap \mathcal{T} \times X$ and the quotient map is the f of Proposition 3.14. Hence

$$\pi \llcorner_{\mathcal{T} \times X \cap \Gamma} = \int_S \pi_y m_\pi(dy), \quad m_\pi = f_\sharp \pi \llcorner_{\mathcal{T} \times X \cap \Gamma}, \ \pi_y \in \mathcal{P}(H(y) \cap \mathcal{T} \times X). \tag{3.2.5}$$

Observe that $m = m_\pi$.

3.3 Regularity of the disintegration

In this Section we consider the translation of Borel sets by the optimal geodesic flow, we introduce the fundamental regularity assumption (Assumption 7) on the measure μ and we show that an immediate consequence is that the set of initial points is negligible and consequently we obtain a disintegration of μ on the whole space. A second consequence is that the disintegration of μ w.r.t. the H has continuous conditions probabilities.

3.3.1 Evolution of Borel sets

Let $A \subset \mathcal{T}_e$ be an analytic set and define for $t \in \mathbb{R}$ the *t-evolution A_t of A* by:

$$A_t := \begin{cases} P_2\big\{(x,y) \in G \cap A \times X : d_N(x,y) = t\big\} & t \geq 0 \\ P_2\big\{(x,y) \in G^{-1} \cap A \times X : d_N(x,y) = t\big\} & t < 0. \end{cases} \tag{3.3.1}$$

It is clear from the definition that if A is analytic, also A_t is analytic . We can show that $t \mapsto \mu(A_t)$ is measurable.

Lemma 3.15. *Let A be analytic. The function $t \mapsto \mu(A_t)$ is $\mathcal{A}$-measurable for $t \in \mathbb{R}$.*

Proof. We divide the proof in three steps.

Step 1. Define the subset of $X \times \mathbb{R}$

$$\hat{A} := \big\{(x,t) : x \in A_t\big\}.$$

Note that

$$\hat{A} = P_{13}\Big\{(x,y,t) \in X \times X \times \mathbb{R}^+ : (x,y) \in G \cap A \times X, d_N(x,y) = t\Big\}$$
$$\cup P_{13}\Big\{(x,y,t) \in X \times X \times \mathbb{R}^- : (x,y) \in G^{-1} \cap A \times X, d_N(x,y) = -t\Big\},$$

hence it is analytic. Clearly $A_t = \hat{A}(t)$.

Step 2. Define the closed set in $\mathcal{P}(X \times [0,1])$

$$\Pi(\mu) := \big\{\pi \in \mathcal{P}(X \times [0,1]) : (P_1)_\sharp(\pi) = \mu\big\}$$

and let $B \subset X \times \mathbb{R} \times [0,1]$ be a Borel set such that $P_{12}(B) = \hat{A}$.

Consider the function

$$\mathbb{R} \times \Pi(\mu) \ni (t,\pi) \mapsto \pi(B(t)).$$

A slight modification of Lemma 4.12 in [8] shows that this function is Borel.

Step 3. Since supremum of Borel function are $\mathcal{A}$-measurable, pag. 134 of [25], the proof is concluded once we show that

$$\mu(A_t) = \mu(\hat{A}(t)) = \sup_{\pi \in \Pi(\mu)} \pi(B(t)).$$

Since $\hat{A}(t) \times [0,1] \supset B(t)$

$$\mu(\hat{A}(t)) = \pi(\hat{A}(t) \times [0,1]) \geq \pi(B(t)).$$

On the other hand from Theorem 5.5, there exists an $\mathcal{A}$-measurable section of the analytic set $B(t)$, so we have $u : \hat{A}(t) \to B(t)$. Clearly for $\pi_u = (\mathbb{I}, u)_\sharp(\mu)$ it holds $\pi_u(B(t)) = \mu(\hat{A}(t))$.

□

The next assumption is the fundamental assumption of the chapter.

Assumption 7 (Non-degeneracy assumption)**.** The measure μ satisfies Assumption 7 if for each analytic set $A \subset \mathcal{T}_e$ there exists a sequence $\{t_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$ and a strictly positive constant C such that $t_n \to 0$ as $n \to +\infty$ and $\mu(A_{t_n}) \geq C\mu(A)$ for every $n \in \mathbb{N}$.

Note that Assumption 7 was already introduced at page 12. Clearly it is enough to verify Assumption 7 for A compact set. An immediate consequence of the Assumption 7 is that the measure μ is concentrated on $\mathcal{T}$.

Proposition 3.16. *If μ satisfies Assumption 7 then*

$$\mu(\mathcal{T}_e \setminus \mathcal{T}) = 0.$$

Proof. Let $A = \mathcal{T}_e \setminus \mathcal{T}$. Suppose by contradiction $\mu(A) > 0$. By the inner regularity there exists $\hat{A} \subset A$ closed with $\mu(\hat{A}) > 0$. By Assumption 7 there exist $C > 0$ and $\{t_n\}_{n\in\mathbb{N}}$ converging to 0 such that $\mu(\hat{A}_{t_n}) \geq C\mu(\hat{A})$.

Define $\hat{A}^\varepsilon := \big\{x : d_N(\hat{A}, x) < \varepsilon\big\}$. Since $\hat{A} \subset A$, for all $n \in \mathbb{N}$ it holds $\hat{A}_{t_n} \cap A = \emptyset$. Moreover for $t_n \leq \varepsilon$ we have $\hat{A}^\varepsilon \supset \hat{A}_{t_n}$. So we have

$$\mu(\hat{A}) = \lim_{\varepsilon \to 0} \mu(\hat{A}^\varepsilon) \geq \mu(\hat{A}) + \mu(\hat{A}_{t_n}) \geq (1 + C)\mu(\hat{A}),$$

that gives the contradiction. □

Once we know that $\mu(\mathcal{T}) = 1$, we can use the Disintegration Theorem 5.3 to write

$$\mu = \int_S \mu_y m(dy), \quad m = f_\sharp \mu, \ \mu_y \in \mathcal{P}(H(y)). \tag{3.3.2}$$

The disintegration is strongly consistent since the quotient map $f : \mathcal{T} \to \mathcal{T}$ is μ-measurable and $(\mathcal{T}, \mathcal{B}(\mathcal{T}))$ is countably generated.

The second consequence of Assumption 7 is that μ_y is continuous, i.e. $\mu_y(\{x\}) = 0$ for all $x \in X$.

Proposition 3.17. *If μ satisfies Assumption 7 then the conditional probabilities μ_y are continuous for m-a.e. $y \in S$.*

Proof. From the regularity of the disintegration and the fact that $m(S) = 1$, we can assume that the map $y \mapsto \mu_y$ is weakly continuous on a compact set $K \subset S$ of comeasure $< \varepsilon$. It is enough to prove the proposition on K.

Step 1. From the continuity of $K \ni y \mapsto \mu_y \in \mathcal{P}(X)$ w.r.t. the weak topology, it follows that the map

$$y \mapsto A(y) := \big\{x \in H(y) : \mu_y(\{x\}) > 0\big\} = \cup_n \big\{x \in H(y) : \mu_y(\{x\}) \geq 2^{-n}\big\}$$

is σ-closed: in fact, if $(y_m, x_m) \to (y, x)$ and $\mu_{y_m}(\{x_m\}) \geq 2^{-n}$, then $\mu_y(\{x\}) \geq 2^{-n}$ by u.s.c. on compact sets. Hence A is Borel, where $A = \{(y, A(y)) : y \in K\}$.

Step 2. The claim is equivalent to $\mu(P_2(A)) = 0$. Suppose by contradiction $\mu(P_2(A)) > 0$. By Lusin Theorem (Theorem 5.8.11 of [25]) A is the countable union of Borel graphs, $A = \cup_n A_n$. Therefore we can take a Borel selection of A just considering one of the Borel graphs, say $\hat{A}$. Since at least one of $P_2(A_n)$ must have positive μ-measure, we can assume $\mu(P_2(\hat{A})) > 0$.

By Assumption 7 $\mu((P_2(\hat{A}))_{t_n}) \geq C\mu(P_2(\hat{A}))$ for some $C > 0$ and $t_n \to 0$. Since $\hat{A}$ is a Borel graph, for every $y \in P_1(\hat{A})$ the set $P_2(\{y\} \times X \cap \hat{A})$ is a singleton. Hence $(P_2(\hat{A}))_{t_n} \cap (P_2(\hat{A}))_{t_m} = \emptyset$. We have a contradiction with the fact that the measure is finite. □

3.4 Solution to the Monge problem

Throughout the section we assume μ to satisfy Assumption 7. It follows from Disintegration Theorem 5.3, Proposition 3.16 and Proposition 3.17 that

$$\mu = \int \mu_y m(dy), \ \pi = \int \pi_y m(dy), \quad \mu_y \text{ continuous}, \ (P_1)_\sharp \pi_y = \mu_y,$$

where $m = f_\sharp \mu$ and $\mu_y \in \mathcal{P}(H(y))$. We write moreover

$$\nu = \int \nu_y m(dy) = \int (P_2)_\sharp \pi_y m(dy).$$

Note that $\pi_y \in \Pi(\mu_y, \nu_y)$ is d_N-cyclically monotone and (since $d_N \llcorner_{H(y) \times H(y)}$ is finite and, from point (a) of Assumption 6, lower semi-continuous) optimal for m-a.e. y. If $\nu(\mathcal{T}) = 1$, then the above formula is the disintegration of ν w.r.t. H.

Theorem 3.18. *Assume that for all $y \in S$ there exists an optimal map T_y from μ_y to ν_y. Then there exists a μ-measurable map $T : X \to X$ such that*

$$\int d_N(x, T(x))\mu(dx) = \int d_N(x, z)\pi(dxdz), \qquad T_\sharp \mu = \nu.$$

Recall $S \subset \mathcal{T}$ introduced in (3.2.3).

Proof. The idea is to use Theorem 5.5.

Step 1. Let $\mathbf{T} \subset S \times \mathcal{P}(X^2)$ be the set: for $y \in S$, $\mathbf{T}_y$ is the family of optimal transference plans in $\Pi(\tilde{\mu}_y, \tilde{\nu}_y)$ concentrated on a graph,

$$\mathbf{T} = \Big\{(y, \pi) \in S \times \mathcal{P}(X^2) : \pi \in \Pi(\mu_y, \nu_y) \text{ optimal}, \exists T : X \to X, \pi(\mathrm{graph}(T)) = 1\Big\}.$$

where for optimal in $\Pi(\mu_y, \nu_y)$ we mean

$$\int d_N \pi = \min_{\pi \in \Pi(\mu_y, \nu_y)} \int d_N \pi.$$

Note that, since π is a Borel measure, in the definition of $\mathbf{T}$, T can be taken Borel. Moreover the y section $\mathbf{T}_y = \mathbf{T} \cap \{y\} \times \mathcal{P}(X^2)$ is not empty.

Step 2. Since the projection is a continuous map, then the set

$$\tilde{\Pi} = \Big\{ (y,\pi) : (P_1)_\sharp \pi = \mu_y, (P_2)_\sharp \pi = \nu_y \Big\}$$

is a Borel subsets of $S \times \mathcal{P}(X^2)$: in fact it is the counter-image of the Borel set graph$((\mu_y,\nu_y)) \subset S \times \mathcal{P}(X)^2$ w.r.t. the weakly continuous map $(y,\pi) \mapsto (y,(P_1)_\sharp\pi,(P_2)_\sharp\pi)$.

Define the Borel function

$$S \times \mathcal{P}(X^2) \ni (y,\pi) \mapsto f(y,\pi) := \begin{cases} \int d_N \pi & \pi \in \Pi(\mu_y,\nu_y) \\ +\infty & \text{otherwise} \end{cases}$$

It follows that $y \mapsto g(y) := \inf_\pi f(y,\pi)$ is an $\mathcal{A}$-function: we can redefine it on a m-negligible set to make it Borel, where m is the quotient measure of μ. Hence the set

$$\tilde{\Pi}^{\mathrm{opt}} = \Big\{ (y,\pi) : \pi \in \tilde{\Pi}(\mu_y,\nu_y), \int d_N \pi \leq g(y) \Big\} = \tilde{\Pi} \cap \Big\{ (y,\pi) : \int d_N \pi \leq g(y) \Big\}$$

is Borel.

Step 3. Now we show that the set of $\pi \in \mathcal{P}(X^2)$ concentrated on a graph is analytic. By Borel Isomorphism Theorem, see [25] page 99, it is enough to prove the same statement for $\pi \in \mathcal{P}([0,1])^2$. Consider the function

$$\mathcal{P}([0,1]^2) \times C_b([0,1],[0,1]) \ni (\pi,\phi) \mapsto h(\pi,\phi) := \pi(\mathrm{graph}(\phi)) \in [0,1].$$

Since graph(ϕ) is compact, h is u.s.c.. Hence the set $B^n = h^{-1}([1-2^{-n},1])$ is closed, so that

$$\mathscr{T} = \bigcap_n P_1(B^n) = \Big\{ \pi : \forall \varepsilon > 0 \ \exists \phi_\varepsilon, \pi(\phi_\varepsilon) > 1 - \varepsilon \Big\}$$

is an analytic set. It is easy to prove that $\pi \in \mathscr{T}$ iff π is concentrated on a graph.

Step 4. It follows that

$$\mathbf{T} = S \times \mathscr{T} \cap \tilde{\Pi}^{\mathrm{opt}}$$

is analytic and by Theorem 5.5 there exists a m-measurable selection $y \mapsto \pi_y \in \mathbf{T}_y$. It is fairly easy to prove that $\int \pi_y m(dy)$ is concentrated on a graph, has the same transference cost of π and belongs to $\Pi(\mu,\nu)$. $\square$

It follows from Theorem 3.18 that it is enough to solve for each $y \in S$ the Monge minimization problem with marginal μ_y and ν_y on the set $H(y)$. In order to solve it, we introduce an assumption on the geometry of the set $H(y)$.

Assumption 8. For a given $y \in S$, $H(y)$ satisfies Assumption 8 if there exist two families of disjoint $\mathcal{A}$-measurable sets $\{K_t\}_{t\in[0,1]}$ and $\{Q_s\}_{s\in[0,1]}$ such that

- $\mu_y(H(y) \setminus \cup_{t\in[0,1]} K_t) = \nu_y(H(y) \setminus \cup_{s\in[0,1]} Q_s) = 0$;
- the associated quotient maps φ_K and φ_Q are respectively μ_y-measurable and ν_y-measurable;
- for $t \leq s$, $K_t \times Q_s \subset G$.

Note that Assumption 8 was already introduced at page 12. In the measurability condition of Assumption 8, the set $[0,1]$ is equipped with the Borel σ-algebra $\mathcal{B}([0,1])$. If $H(y)$ satisfies Assumption 8 we can disintegrate the marginal measures μ_y and ν_y respectively w.r.t. the family $\{K_t\}$ and $\{Q_s\}$:

$$\mu_y = \int \mu_{y,t} m_{\mu_y}(dt), \quad \nu_y = \int \nu_{y,t} m_{\nu_y}(dt)$$

where $m_{\mu_y} = \varphi_{K\,\sharp}\mu_y$, $m_{\nu_y} = \varphi_{Q\,\sharp}\nu_y$ and the disintegrations are strongly consistent.

Proposition 3.19. *Suppose that $H(y)$ satisfies Assumption 8 and that the following conditions hold true:*

- m_{μ_y} *is continuous;*
- $\mu_{y,t}$ *is continuous for* m_{μ_y}*-a.e.* $t \in [0,1]$*;*
- $m_{\mu_y}([0,t]) \geq m_{\nu_y}([0,t])$ *for* m_{μ_y}*-a.e.* $t \in [0,1]$.

Then there exists a d_N-cyclically monotone μ_y-measurable map T_y such that $T_{y\,\sharp}\mu_y = \nu_y$ and

$$\int d_N(x, T_y(x))\mu_y(dx) = \int d_N(x,z)\pi_y(dxdz).$$

Proof. Step 1. Since m_{μ_y} is continuous and $m_{\mu_y}([0,t]) \geq m_{\nu_y}([0,t])$, there exists an increasing map $\psi : [0,1] \to [0,1]$ such that $\psi_\sharp m_{\mu_y} = m_{\nu_y}$.

Moreover, since for m_{μ_y}-a.e. $t \in [0,1]$ $\mu_{y,t}$ is continuous, there exists a Borel map $T_t : K_t \to Q_{\psi(t)}$ such that $T_{t\,\sharp}\mu_{y,t} = \nu_{y,\psi(t)}$ for m_{μ_y}-a.e. $t \in [0,1]$. Since $\psi(t) \geq t$ the map T_t is d_N-cyclically monotone, hence optimal between $\mu_{y,t}$ and $\nu_{y,t}$.

Step 2. Reasoning as in the proof of Theorem 3.18, one can prove the existence of a μ_y-measurable map $T_y : H(y) \to H(y)$ that is the gluing of all the maps T_t constructed in *Step 1.*. Hence there exists a μ_y-measurable map $T_y : H(y) \to H(y)$ such that $T_{y\,\sharp}\mu_{y,t} = \nu_{y,\psi(t)}$. It follows from Assumption 8 that $\mathrm{graph}(T_y) \subset G$, hence T_y is d_N-cyclically monotone and

$$T_\sharp\mu_y = \int T_\sharp\mu_{y,t} m_{\mu_y}(dt) = \int \nu_{y,\psi(t)} m_{\mu_y}(dt) = \int \nu_{y,t}(\psi_\sharp m_{\mu_y})(dt) = \nu_y.$$

□

The next corollary follows straightforwardly and it sums up all the results.

Corollary 3.20. *Let $\pi \in \Pi(\mu,\nu)$ be concentrated on a d_N-cyclically monotone set Γ satisfying Assumption 6. Assume that μ satisfies Assumption 7 and for m-a.e. $y \in S$ the set $H(y)$ satisfies Assumption 8. If for m-a.e. $y \in S$ the hypothesis of Proposition 3.19 are verified, then there exists an Borel map $T : X \to X$ such that*

$$\int d_N(x, T(x))\mu(dx) = \int d_N(x,z)\pi(dxdz), \qquad T_\sharp\mu = \nu.$$

If π is also optimal, then T solves the Monge minimization problem.

Let us summarize the theoretical results obtained so far. Let $\pi \in \Pi(\mu,\nu)$ be a d_N-cyclically monotone transference plan concentrated on a set Γ. Consider the corresponding family of chain of transport rays and assume that Γ satisfies Assumption 6. Then the partition induced by H permits to obtain a strongly consistent disintegration formula of μ, ν and π holds. If μ satisfies Assumption 7 then the set of initial points is μ-negligible and the conditional probabilities μ_y are continuous.

Since the geometry of $H(y)$ can be wild, we need another assumption to build a d_N-monotone transference map between μ_y and ν_y. If $H(y)$ satisfies Assumption 8 we can perform another disintegration and, under additional regularity of the conditional probabilities of μ_y and of the quotient measure of μ_y, we prove the existence of a d_N-monotone transference map between μ_y and ν_y. Applying the same reasoning for m-a.e. y we prove the existence of a transport map T between μ and ν that has the same transference cost of the given d_N-cyclically monotone plan π.

3.4.1 Example

We conclude this Section with the analysis of a particular case in which the set $H(y)$ satisfies Assumption 8. The hypothesis of Proposition 3.19 and Assumption 8 were partially inspired by this example. What

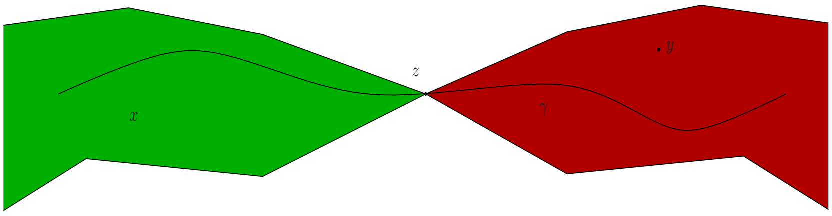

Figure 3.1: The hourglass set $K(z)$.

follows will be useful in the next Section, however, since it is not only related to what will be proved in Section 3.5, we have decided to present it here.

Fix the following notation: a continuous curve $\gamma : [0,1] \to X$ is *increasing* if for $t, s \in [0,1]$

$$t \leq s \Longrightarrow (\gamma(t), \gamma(s)) \in G$$

Definition 3.21 (Hourglass sets)**.** For $z \in X$ define the *hourglass set*

$$K(z) := \Big\{(x,y) \in X \times X : (x,z),(z,y) \in G\Big\}.$$

Assume that there exists an increasing curve γ such that

$$H(y) \times X \cap \Gamma \subset \bigcup_{t \in [0,1]} K(\gamma(t)) \cap \Gamma.$$

Note that this assumption is equivalent to request that on each chain of transport rays the branching structures can appear only along an increasing curve γ.

Then $H(y)$ satisfies Assumption 8. Indeed first notice that $K(z)$ is analytic, then define the family of sets

$$K_t := G^{-1}(\gamma(t)) \setminus \bigcap_{s<t} G^{-1}(\gamma(s)), \qquad Q_t := G(\gamma(t)) \setminus \bigcap_{t<s} G(\gamma(s)).$$

Since γ is increasing, K_t and Q_s are $\mathcal{A}$-measurable and the quotient maps are $\mathcal{A}$-measurable: let $[a,b] \subset [0,1]$

$$\varphi_K^{-1}([a,b]) = \bigcup_{t \in [a,b]} K_t = G^{-1}(\gamma(b)) \setminus G^{-1}(\gamma(a)) \cup K_a \in \mathcal{A}$$

and the same calculation holds true for φ_Q. From the increasing property of γ it follows that $K_t \times Q_s \in G$ for every $0 \leq t \leq s \leq 1$. Again from the increasing property of γ it follows that $m_{\mu_y}([0,t]) \geq m_{\nu_y}([0,t])$.

3.5 The Obstacle problem

Throughout this section $|\cdot|$ will be the euclidean distance of $\mathbb{R}^d$.

Let $C \subset \mathbb{R}^d$ be an open convex set such that $M := \partial C$ is a smooth compact sub-manifold of $\mathbb{R}^d$ of dimension $d-1$. Let $X := \mathbb{R}^d \setminus C$. Clearly X endowed with the euclidean topology is a Polish space.

Consider the following geodesic distance: $d_M : X \times X \to [0, +\infty]$:

$$d_M(x,y) := \inf\{L(\gamma) : \gamma \in \mathrm{Lip}([0,1], X), \gamma(0) = x, \gamma(1) = y\}, \tag{3.5.1}$$

where L is the standard euclidean arc-length: $L(\gamma) = \int |\dot{\gamma}|$. Hence M can be seen as an obstacle for geodesics connecting points in X. Note that any minimizing sequence has uniformly bounded Lipschitz constant, therefore in the definition of d_M we can substitute inf with min. Hence d_M is a geodesic distance on X.

We will show that given $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ with $\mu \ll \mathcal{L}^d$, the Monge minimization problem with geodesic cost d_M admits a solution.

From now on we will assume that $\mu \ll \mathcal{L}^d$ and all the sets and structures introduced during the chapter will be referred to this Monge problem.

The strategy to solve the Monge minimization problem is the one used in Section 3.4: build an optimal map on each equivalence class $H(y)$ and then use Theorem 3.18. To prove the existence these optimal maps we will show that the geometry of the chain of transport rays $H(y)$ is the one presented in Example 3.4.1 and that the hypothesis of Proposition 3.19 are satisfied.

Lemma 3.22. *The distance d_M is a continuous map.*

Proof. Step 1. Let $\{x_n\}_{n\in\mathbb{N}}, \{y_n\}_{n\in\mathbb{N}} \in X$ such that $|x_n - x| \to 0$ $|y_n - y| \to 0$. Since the boundary of X is a smooth manifold, for every $n \in \mathbb{N}$ there exist curves $\gamma_{1,n}, \gamma_{2,n} \in \mathrm{Lip}([0,1], X)$ such that

- $\gamma_{1,n}(0) = x, \gamma_{1,n}(1) = x_n$;
- $\gamma_{2,n}(0) = y, \gamma_{2,n}(1) = y_n$;
- $L(\gamma_{i,n}) \to 0$ as $n \to +\infty$, for $i = 1, 2$.

Consider $\gamma_n \in \mathrm{Lip}([0,1], X)$ such that $\gamma_n(0) = x_n, \gamma_n(1) = y_n$ and $L(\gamma_n) \leq d_M(x_n, y_n) + 2^{-n}$. Gluing $\gamma_{1,n}$ and $\gamma_{2,n}$ to γ_n it follows

$$d_M(x, y) \leq d_M(x_n, y_n) + 2^{-n} + L(\gamma_{1,n}) + L(\gamma_{2,n}).$$

Hence d_M is l.s.c..

Step 2. Taking a minimizing sequence of admissible curves for $d_M(x, y)$ and gluing them with $\gamma_{i,n}$ as in *Step 1.*, it is fairly easy to prove that d_M is u.s.c. and therefore continuous. □

As a corollary we have the existence of an optimal transference plan π. Hence from now on π will be an optimal transference plan and all the structures defined during the chapter starting from a generic d_N-cyclically monotone plan, are referred to it. Moreover there exists $\varphi \in \mathrm{Lip}_{d_M}(X, \mathbb{R})$ such that $\Gamma = \Gamma' = G = \{(x, y) \in X \times X : \varphi(x) - \varphi(y) = d_M(x, y)\}$. Note that Γ is closed.

The next result shows that the sets $H(y)$ have the structure of Example 3.4.1. The convex assumption on the obstacle is fundamental: each transport rays is composed by a straight line, a geodesic on M where branching structures are allowed and again a straight line.

Lemma 3.23. *For all $y \in S$, $H(y)$ has the geometry of Example 3.4.1: there exists an increasing curve $\gamma_y : [0,1] \to X$ such that*

$$H(y) \times X \cap \Gamma \subset \bigcup_{t\in[0,1]} K(\gamma_y(t)) \cap \Gamma.$$

Proof. Since due to convexity and smoothness of the obstacle, the geodesics of d_M are smooth and composed by a first straight line, a geodesic of the manifold and a final straight line, a branching structure can appear only on the manifold M. If $H(y) \neq R(y)$, consider the following sets:

$$Z := \bigcap_{z\in H(y)\cap M} G^{-1}(z) \cap M, \quad W := \bigcap_{z\in H(y)\cap M} G(z) \cap M.$$

By d_M-monotonicity, smoothness and convexity of M, for all $z \in H(y) \cap M$ the set $G^{-1}(z) \cap M$ is always contained in the same geodesic of M. Using the compactness of M, $Z = \{z\}$ and $W = \{w\}$ and $(z, w) \in G$. Consider the unique increasing geodesic $\gamma_y \in \gamma_{[z,w]}$ such that $\gamma_y = G(z) \cap G^{-1}(w)$. Hence

$$H(y) \times X \cap \Gamma \subset \bigcup_{t\in[0,1]} K(\gamma_y(t)) \cap \Gamma.$$

□

Remark 3.24. From Lemma 3.22 and Lemma 3.23 it follows that, for any transference plan π, the set of chain of transport rays H satisfies Assumption 6.

Indeed consider $H(y)$ and the corresponding geodesic γ_y from Lemma 3.23. Then take any sequence $x_n \in H(y)$ such that $|x_n - x| \to 0$ as $n \to +\infty$. Note that there exist $s_n \in [0,1]$ and $t_n \in \mathbb{R}$ such that $x_n = \gamma_y(s_n) + t_n \nabla\gamma_y(s_n)$. Possibly passing to subsequences, $s_n \to s$, $t_n \to t$ with $x = \gamma_y(s) + t\nabla\gamma_y(s)$. Since $(\gamma_y(s_n), x_n) \in G$ and G is closed it follows that $(\gamma_y(s), x) \in G$. From $(y, \gamma_y(s)) \in R$ follows $x \in H(y)$. Hence point (a) of Assumption 6 holds true.

Point (b) of Assumption 6 follows directly from the continuity of d_N.

In the following Lemma we prove that the problem can be reduced to the equivalence classes $H(y)$. We use the following notation: the quotient map induced by H will be denoted by f^y and the corresponding quotient measure $f^y_\sharp \mu$ by m_H.

Lemma 3.25. *The μ-measure of the set of initial points is zero, hence*

$$\mu = \int \mu_y m_H(dy).$$

Moreover μ_y is continuous for m_H-a.e. y.

Proof. Step 1. Since $\mu \ll \mathcal{L}^d$, it is enough to prove that the set of initial points is $\mathcal{L}^d$-negligible and that the disintegration w.r.t. H of $\mathcal{L}^d$ restricted to any compact set has continuous conditional probabilities. Indeed if $\mathcal{L}^d \llcorner_K = \int \eta_y m_{\mathcal{L}^d}(dy)$ and $\mu = \rho\mathcal{L}^d$ then $m_\mu \ll m_{\mathcal{L}^d}$ and

$$\mu\llcorner_K = \int \rho\eta_y m_{\mathcal{L}^d}(dy) = \int \rho \frac{dm_{\mathcal{L}^d}}{dm_\mu} \eta_y m_\mu(dy),$$

where m_μ is the quotient measure of $\mu\llcorner_K$. It follows that the continuity of η_y implies the continuity of conditional probabilities of μ. Hence the claim is to prove that $\mathcal{L}^d$ satisfies Assumption 7.

Step 2. Let $K \subset X$ be any compact set with $\mathcal{L}^d(K) > 0$. Possibly intersecting K with $B_r(x)$ for some $x \in \mathbb{R}^d \setminus C$ and $r > 0$, we can assume w.l.o.g. that $K \subset B_\varepsilon(x)$ and $B_{2\varepsilon}(x) \cap M = \emptyset$. Since $d_M \geq d$, $K_t \subset B_{2\varepsilon}(x)$ for all $t \leq \varepsilon$. Since $d_M = |\cdot|$ in $B_{2\varepsilon}(x)$, it follows that inside $B_{2\varepsilon}(x) \times B_{2\varepsilon}(x)$ d_M-cyclically monotonicity is equivalent to $|\cdot|$-cyclically monotonicity. It follows that the set $H \cap G(K) \times G^{-1}(K_\varepsilon)$ is $|\cdot|$-cyclically monotone.

Step 3. The following is proved in Chapter 1: consider a metric measure space (X, d, m) with d non-branching geodesic distance, $m \in \mathcal{P}(X)$ and assume that (X, d, m) satisfies $MCP(K, N)$. Let Γ be a d-cyclically monotone set and consider the evolution of sets induced by Γ, then m satisfies Assumption 7 w.r.t. this evolution of sets.

Since $B_{2\varepsilon}(x)$ is a convex set, it follows that $(B_{2\varepsilon}(x), |\cdot|, \mathcal{L}^d)$ satisfies $MCP(0, d)$. Therefore $\mathcal{L}^d$ satisfies Assumption 7 w.r.t. the evolution of sets induced by $H \cap G(K) \times G^{-1}(K_\varepsilon)$. The claim follows. □

Hence we can assume w.l.o.g. that $\mu(G^{-1}(M)) = \nu(G(M)) = 1$: if $H(y)$ do not intersect the obstacle, it is a straight line and the marginal μ_y is continuous. Since the existence of an optimal transport map on a straight line with first marginal continuous is a standard fact in optimal transportation, the reduction follows.

Recall the two family of sets introduced in Example 3.4.1:

$$K_{y,t} := G^{-1}(\gamma_y(t)) \setminus \bigcap_{s<t} G^{-1}(\gamma_y(s)), \qquad Q_{y,t} := G(\gamma_y(t)) \setminus \bigcap_{t<s} G(\gamma_y(s)).$$

It follows from Lemma 3.23 and Example 3.4.1 that

$$\mu_y = \int \mu_{y,t} m_{\mu_y}(dt), \quad \nu_y = \int \nu_{y,t} m_{\nu_y}(dt).$$

with $\mu_{y,t}(K_{y,t}) = \nu_{y,t}(Q_{y,t}) = 1$. Moreover using the increasing curve γ_y, we can assume that $m_{\mu_y} \in \mathcal{P}(M)$, indeed

$$\mu_y = \int_{[0,1]} \mu_{y,t} m_{\mu_y}(dt) = \int_{\gamma_y([0,1])} \mu_{y,\gamma_y^{-1}(z)}(\gamma_{y\,\sharp} m_{\mu_y})(dz). \tag{3.5.2}$$

And the same calculation holds true for ν_y and m_{ν_y}. Therefore in the following

$$\mu_y = \int_M \mu_{y,z} m_{\mu_y}(dz), \quad \nu_y = \int_M \nu_{y,z} m_{\nu_y}(dz) \tag{3.5.3}$$

with $\mu_{y,z}(K_{y,\gamma_y^{-1}(z)}) = \nu_{y,z}(Q_{y,\gamma_y^{-1}(z)}) = 1$ and $m_{\mu_y}(\gamma_y([0,1])) = m_{\nu_y}(\gamma_y([0,1])) = 1$.

Moreover w.l.o.g. we can assume that $S = f^y(\mathbb{R}^d) \subset M$, in particular we can assume that for all $y \in S$ there exists $t(y) \in [0,1]$ such that $y = \gamma_y(t(y))$.

According to Proposition 3.19, to obtain the existence of an optimal map on $H(y)$ it is enough to prove that m_{μ_y} is continuous and $\mu_{y,z}$ is continuous for m_{μ_y}-a.e. $z \in M$. Recall that $m_{\mu_y}(\gamma_y([0,t])) \geq m_{\nu_y}(\gamma_y([0,t]))$ is a straightforward consequence of the increasing property of γ_y.

Remark 3.26. Consider the following $\mathcal{A}$-measurable map:

$$G^{-1}(M) \setminus (a(M) \cap M) \ni w \mapsto f^M(w) := \operatorname{Argmin}\{d(z,w) : z \in M \cap G(w)\} \in M.$$

Consider the measure $m := f^M_\sharp \mu \in \mathcal{P}(M)$. Observing that $f^M(H(y)) = \gamma_y([0,1])$, it follows that the support of m is partitioned by a d_M-cyclically monotone equivalence relation:

$$m\Big(\bigcup_{y\in S} \gamma_y([0,1])\Big) = 1, \qquad \bigcup_{y\in S} \gamma_y([0,1]) \times \bigcup_{y\in S} \gamma_y([0,1]) \cap G \ \text{ is } d_M\text{-cyclically monotone}$$

Moreover f^y is a quotient map also for this equivalence relation. Note that $f^y_\sharp m = m_H$: consider $I \subset S$

$$\begin{aligned}(f^y_\sharp m)(I) &= m\Big(\bigcup_{y\in I} \gamma_y([0,1])\Big) = \mu\Big(G^{-1}(\bigcup_{y\in I} \gamma_y([0,1]))\Big) \\ &= \mu\Big(\bigcup_{y\in I} H(y)\Big) = (f^y_\sharp \mu)(I) = m_H(I).\end{aligned}$$

It follows that

$$m = \int_S (f^M_\sharp \mu_y) m_H(dy)$$

and from (3.5.3) $f^M_\sharp \mu_y = m_{\mu_y}$. Hence the final disintegration formula for m is the following one:

$$m = \int_S m_{\mu_y} m_H(dy) \tag{3.5.4}$$

Proposition 3.27. *The measure m is absolutely continuous w.r.t. the Hausdorff measure $\mathcal{H}^{d-1}$ restricted to M.*

Proof. Recall that $\varphi \in \mathrm{Lip}_{d_M}(\mathbb{R}^d)$ is the potential associated to Γ and consider the following set

$$M_2 := P_1\Big(\{(x,y) \in M \times M : |\varphi(x) - \varphi(y)| = d_M(x,y)\} \setminus \{x = y\}\Big).$$

Step 1. Define the following map: $M_2 \ni w \mapsto \Xi(w) := \operatorname{Argmin}\{\varphi(w) - \varphi(z) : z \in M_2\}$. Then the function φ is a potential for the Monge minimization problem on M with cost the geodesic distance, that coincides with d_M, with first marginal m and as second marginal $\Xi_\sharp m$.

It follows from Proposition 15 of [20] that $\nabla\varphi$ is a Lipschitz function: for all $x, y \in M_2$

$$|\nabla\varphi(x) - \nabla\varphi(y)| \leq L d_M(x, y).$$

In [20] the Lipschitz constant L is uniform for x, y belonging to sets uniformly far from the starting and ending points of the geodesics on M of the transport set. Since in our setting the geodesics on M do not intersect, L is uniform on the whole M. Moreover note that if $z = \gamma_y(t)$, then

$$\nabla\varphi(z) = -\frac{\dot{\gamma}_y(t)}{|\dot{\gamma}_y(t)|}.$$

Step 2. For $t \geq 0$, define the following map

$$M_2 \ni x \mapsto \psi_t(x) := x + \nabla\varphi(x)t.$$

Possibly restricting ψ_t to a subset of M of points coming from transport rays of uniformly positive length, since $t \mapsto \psi_t(x)$ is a parametrization of the transport ray touching M in x, by d_M-cyclical monotonicity of Γ, we can assume that ψ_t is injective. Moreover ψ_t is bi-Lipschitz, provided t is small enough: indeed

$$|x + \nabla\varphi(x)t - y - \nabla\varphi(y)t| \geq |x - y|(1 - Lt).$$

It follows that

$$M_2 \times [-\delta, \delta] \ni (x, s) \mapsto \psi(x, s) := x + \nabla\varphi(x)(t + s)$$

is bi-Lipschitz and injective provided $\delta \leq 1/L + t$. Hence the Jacobian determinant of φ, $Jd\varphi$, is uniformly positive.

Step 3. Consider the following set

$$B := \{x \in \mathbb{R}^d : t - \delta \leq d(M, x) \leq t + \delta\} \cap G^{-1}(M)$$

where d is the euclidean distance. Clearly B is the range of ψ and $\mathcal{L}^d(B) > 0$. Since M is a smooth manifold, we can pass to local charts: let $U_\alpha \subset \mathbb{R}^{d-1}$ be an open set and $h_\alpha : U_\alpha \to M$ the corresponding parametrization map. The map

$$U_\alpha \times [-\delta, \delta] \ni (x, s) \mapsto \psi_\alpha(x, s) := \psi(h_\alpha(x), s)$$

is a bi-Lipschitz parametrization of the set $B_\alpha := B \cap G^{-1}(h_\alpha(U_\alpha))$.

It follows directly from the Area Formula, see for example [1], that

$$\mathcal{L}^d \llcorner_{B_\alpha} = \psi_{\alpha\,\sharp}\Big(Jd\psi_\alpha(\mathcal{L}^{d-1} \times dt)\llcorner_{U_\alpha \times [-\delta,\delta]}\Big),$$

hence $f^M_\sharp \mathcal{L}^d \llcorner_{B_\alpha} \ll \mathcal{H}^{d-1} \llcorner_M$. Since B can be covered with a finite number of B_α and $\mathcal{L}^d \llcorner_{B_\alpha}$ is equivalent to m, the claim follows. □

Recall the following result. Let (M, g) be a n-dimensional compact Riemannian manifold, let d_M be the geodesic distance induced by g and η the volume measure. Then the disintegration of η w.r.t. any d_M-cyclically monotone set is strongly consistent and the conditional probabilities are continuous. This result is proved in [9], Theorem 9.5, in the more general setting of metric measure space satisfying the measure contraction property.

Corollary 3.28. *For m_H-a.s. $y \in S$, the quotient measure m_{μ_y} is continuous.*

Proof. We have proved in Remark 3.26 that the measures m_{μ_y} are the conditional probabilities of the disintegration of m w.r.t. the equivalence relation given by the membership to geodesics γ_y and m_H is the corresponding quotient measure. Hence the claim follows directly from Theorem 9.5 of [9] and Proposition 3.27. □

Proposition 3.29. *For m_H-a.e. $y \in S$, the measures $\mu_{y,z}$ are continuous for m_{μ_y}-a.e. $z \in M$.*

Proof. Recall that $f^M_\sharp \mu = m$.

Step 1. The measure μ can be disintegrated w.r.t. the partition given by the family of pre-images of the $\mathcal{A}$-measurable map f^M: $\{(f^M)^{-1}(p)\}_{p \in f^M(\mathbb{R}^d)}$. Clearly f^M is a possible quotient map, hence

$$\mu = \int \mu_z m(dz), \tag{3.5.5}$$

The set $G^{-1}(M) \setminus a(M) \times G^{-1}(M) \setminus a(M) \cap G$ is $|\cdot|$-cyclically monotone and $\mu \ll \mathcal{L}^d$, hence it follows that for m-a.e. $z \in f^M(\mathbb{R}^d)$, μ_z is continuous.

Step 2. From Lemma 3.25 $\mu = \int \mu_y m_H(dy)$, therefore

$$m = f^M_\sharp \mu = \int (f^M_\sharp \mu_y) m_H(dy),$$

hence using (3.5.5) and the uniqueness of the disintegration

$$\mu = \int \bigg(\int \mu_z (f^M_\sharp \mu_y)(dz) \bigg) m_H(dy), \qquad \mu_y = \int \mu_z (f^M_\sharp \mu_y)(dz),$$

where the last equality holds true for m_H-a.e. $y \in S$. Hence for m_H-a.e. $y \in S$ the measures $\mu_{y,z}$ are continuous for m_{μ_y}-a.e. $z \in M$. □

Finally we can prove the existence of an optimal map for the Monge minimization problem with obstacle.

Theorem 3.30. *There exists a solution for the Monge minimization problem with cost d_M and marginal μ, ν with $\mu \ll \mathcal{L}^d$.*

Proof. From Lemma 3.25 it follows that μ can be disintegrated w.r.t. the equivalence relation H. From Theorem 3.18 it follows that to prove the claim it is enough to prove the existence of an optimal map on each equivalence class $H(y)$. Hence we restrict the analysis to the classes $H(y)$ such that $H(y) \neq R(y)$ and for them we proved in Lemma 3.23 that Assumption 8 holds true. In Proposition 3.27, Corollary 3.28 and Proposition 3.29 we proved that for m_H-a.e. $y \in S$ the measures m_{μ_y} and $\mu_{y,z}$ verify the hypothesis of Proposition 3.19. Therefore the claim follows. □

Chapter 4

Local curvature-dimension condition implies measure-contraction property

We present the main steps of this chapter.

Section 4.1 contains some classical preliminary results. The references are [7, 27]. Section 4.1 starts with the definitions of metric measure space, of length space and of the non-branching condition. In Section 4.1.1 we first recall the main ingredients involved in $\mathsf{CD}(K,N)$: the Rényi entropy (4.1.1) and the volume distortion coefficient (4.1.2). Then the definition of $\mathsf{CD}(K,N)$, $\mathsf{CD}_{loc}(K,N)$ and $\mathsf{MCP}(K,N)$ are given.

In Section 4.2, given $x_0 \in M$, for every fixed $\bar{R} > 0$, we write m in "polar coordinates": considering the partition of $B_{\bar{R}}(x_0)$ given by $\{S_r(x_0)\}_{r \leq \bar{R}}$ where $S_r(x_0) = \partial B_{\bar{R}}(x_0)$, Disintegration Theorem implies

$$m \llcorner_{\bar{B}_{\bar{R}}(x_0)} = \int_{[0,\bar{R}]} \mathcal{S}_p dp, \qquad \mathcal{S}_p(M \setminus S_p(x_0)) = 0.$$

Then we study the geodesic curve $\mathcal{G}_{\bar{R}} : [0,\bar{R}] \to \mathcal{M}_+(M)$ such that $\mathcal{G}_{\bar{R}}(0) = \delta_{x_0} \|\mathcal{S}_{\bar{R}}\|$ and $\mathcal{G}_{\bar{R}}(\bar{R}) = \mathcal{S}_{\bar{R}}$. We prove (Lemma 4.9) that $\mathcal{G}_{\bar{R}}(p) = h_{\bar{R}}(\cdot,p)\mathcal{S}_p$ and (Lemma 4.11) the map $[0,\bar{R}] \ni p \mapsto h_{\bar{R}}(y,p)$ is locally Lipschitz continuous.

Since our aim is to understand the behavior of the geodesic curve $p \mapsto \mathcal{G}_{\bar{R}}(p)$, it is now sufficient to understand the behavior of its density. In Section 4.3 we prove that the map $[0,\bar{R}] \ni p \mapsto h_{\bar{R}}(y,p)$ satisfies a sort of $\mathsf{CD}^*_{loc}(K,N-1)$ that can be written in the following way: for every $0 < r < \bar{R}$ there exists ε_r such that (Theorem 4.12) if $R - r < \varepsilon_r$, then the following holds true

$$h_{\bar{R}}(r + t(R-r)))^{-\frac{1}{N-1}} \geq h_{\bar{R}}(r)^{-\frac{1}{N-1}} \sigma^{(1-t)}_{K,N-1}(R-r) + h_{\bar{R}}(R)^{-\frac{1}{N-1}} \sigma^{(t)}_{K,N-1}(R-r), \tag{4.0.1}$$

for every $t \in [0,1]$.

In Section 4.4 we show that (Theorem 4.14) the above expression (4.0.1) satisfies a local to global property and therefore the restriction $R - r < \varepsilon_r$ can be removed and a global version of (4.0.1) is obtained.

In Section 4.5, exploiting the expression of m is polar coordinate and the informations obtained in Section 4.4 on the $(N-1)$-dimensional evolution, we prove that (Proposition 4.15 and Theorem 4.16) the non-branching metric measure space (M,d,m) satisfies the condition $\mathsf{MCP}(K,N)$.

This chapter is based on the joint work with Karl-Theodor Sturm [17].

4.1 Preliminaries

Let (M,d) be a metric space. The length $\mathsf{L}(\gamma)$ of a continuous curve $\gamma:[0,1]\to M$ is defined as

$$\mathsf{L}(\gamma):=\sup\sum_{k=1}^{n}d(\gamma(t_{k-1}),\gamma(t_k))$$

where the supremum runs over $n\in\mathbb{N}$ and over all partitions $0=t_0<t_1<\cdots<t_n=1$. Clearly $\mathsf{L}(\gamma)\geq d(\gamma(0),\gamma(1))$. The curve is called *geodesic* if and only if $\mathsf{L}(\gamma)=d(\gamma(0),\gamma(1))$. In this case we always assume that γ has constant speed, i.e. $\mathsf{L}(\gamma\llcorner_{[s,t]})=|s-t|\mathsf{L}(\gamma)=|s-t|d(\gamma(0),\gamma(1))$ for every $0\leq s\leq t\leq 1$.

With $\mathcal{G}(M)$ we denote the space of geodesic $\gamma:[0,1]\to M$ in M, regarded as subset of $Lip([0,1],M)$ of Lipschitz functions equipped with the topology of uniform convergence.

(M,d) is said a *length space* if and only if for all $x,y\in M$,

$$d(x,y)=\inf\mathsf{L}(\gamma)$$

where the infimum runs over all continuous curves connecting x to y. It is said to be a *geodesic space* if and only if every $x,y\in M$ are connected by a geodesic.

Definition 4.1. A geodesic space (M,d) is *non-branching* if and only if for all $r\geq 0$ and $x,y\in M$ such that $d(x,y)=r/2$ the set

$$\{z\in M:d(x,z)=r\}\cap\{z\in M:d(y,z)=r/2\}$$

is a singleton.

A *metric measure space* will always be a triple (M,d,m) where (M,d) is a complete separable metric space and m is a locally finite measure (i.e. $m(Br(x))<\infty$ for all $x\in M$ and all sufficiently small $r>$0) on M equipped with its Borel σ-algebra. We exclude the case $m(M)=0$. A *non-branching* metric measure space will be a metric measure space (M,d,m) such that (M,d) is a non-branching geodesic space.

Throughout the following we will use the notation:

$$S_p(z)=\{x:d(x,z)=p\},\qquad B_p(z)=\{x:d(x,z)<p\}.$$

4.1.1 Geometry of Metric measure spaces

$\mathcal{P}_2(M,d)$ denotes the L^2-Wasserstein space of probability measures on M and d_W the corresponding L^2-Wasserstein distance. The subspace of m-absolutely continuous measures is denoted by $P_2(M,d,m)$. A point z will be called t-intermediate point of points x and y if $d(x,z)=td(x,y)$ and $d(z,y)=(1-t)d(x,y)$.

The following are well-known results in optimal transportation and are valid for general metric measure spaces.

Lemma 4.2. *Let (M,d,m) be a metric measure space. For each geodesic $\Gamma:[0,1]\to\mathcal{P}_2(M)$ there exists a probability measure Ξ on $\mathcal{G}(M)$ such that*

- $e_{t\,\sharp}\Xi=\Gamma(t)$ *for all $t\in[0,1]$;*
- *for each pair (s,t) the transference plan $(\gamma_s,\gamma_t)_\sharp\Xi$ is an optimal coupling.*

The curvature-dimension condition $\mathsf{CD}(K,N)$ is defined in terms of convexity properties of the lower semi-continuous Rényi entropy functional

$$\mathcal{E}_N(\mu|m):=-\int_M\varrho^{-1/N}(x)\mu(dx)\tag{4.1.1}$$

on $P_2(M,d)$ where ϱ denotes the density of the absolutely continuous part μ^c in the Lebesgue decomposition $\mu = \mu^c + \mu^s = \varrho m + \mu^s$.

Given two numbers $K, N \in \mathbb{R}$ with $N \geq 1$, we put for $(t,\theta) \in [0,1] \times \mathbb{R}_+$,

$$\tau_{K,N}^{(t)}(\theta) := \begin{cases} \infty, & \text{if } K\theta^2 \geq (N-1)\pi^2, \\ t^{1/N}\left(\dfrac{\sin(t\theta\sqrt{K/(N-1)})}{\sin(\theta\sqrt{K/(N-1)})}\right)^{1-1/N} & \text{if } K\theta^2 \leq (N-1)\pi^2, \\ t & \text{if } K\theta^2 < 0 \text{ or} \\ & \text{if } K\theta^2 = 0 \text{ and } N = 1, \\ t^{1/N}\left(\dfrac{\sinh(t\theta\sqrt{-K/(N-1)})}{\sinh(\theta\sqrt{-K/(N-1)})}\right)^{1-1/N} & \text{if } K\theta^2 \leq 0 \text{ and } N > 1. \end{cases} \tag{4.1.2}$$

That is, $\tau_{K,N}^{(t)}(\theta) := t^{1/N}\sigma_{K,N-1}^{(t)}(\theta)^{(N-1)/N}$ where

$$\sigma_{K,N}^{(t)}(\theta) = \frac{\sin(t\theta\sqrt{K/N})}{\sin(\theta\sqrt{K/N})},$$

if $0 < K\theta^2 < N\pi^2$ and with appropriate interpretation otherwise. Moreover we put

$$\varsigma_{K,N}^{(t)}(\theta) := \tau_{K,N}^{(t)}(\theta)^N.$$

The coefficients $\tau_{K,N}^{(t)}(\theta), \sigma_{K,N}^{(t)}(\theta)$ and $\varsigma_{K,N}^{(t)}(\theta)$ are all volume distortion coefficients depending on the curvature K and on the dimension N.

Definition 4.3 (Curvature-Dimension condition)**.** Let two number $K, N \in \mathbb{R}$ with $N \geq 1$ be given. We say that (M,d,m) satisfies the curvature-dimension condition - denoted by $\mathsf{CD}(K,N)$ - if and only if for each pair $\nu_0, \nu_1 \in \mathcal{P}_2(M,d,m)$ there exists an optimal coupling π of $\nu_0 = \varrho_0 m$ and $\nu_1 = \varrho_1 m$, and a geodesic $\Gamma : [0,1] \to \mathcal{P}_2(M,d,m)$ connecting ν_0 and ν_1 with

$$\begin{aligned} \mathcal{E}_{N'}(\Gamma(t)|m) \leq -\int_{M\times M} \Big[&\tau_{K,N'}^{(1-t)}(d(x_0,x_1))\varrho_0^{-1/N'}(x_0) \\ &+ \tau_{K,N'}^{(t)}(d(x_0,x_1))\varrho_1^{-1/N'}(x_1)\Big]\pi(dx_0dx_1), \end{aligned} \tag{4.1.3}$$

for all $t \in [0,1]$ and all $N' \geq N$.

We recall also the definition of the reduced curvature-dimension condition $\mathsf{CD}^*(K,N)$ introduced in [7].

Definition 4.4 (Reduced Curvature-Dimension condition)**.** Let two number $K, N \in \mathbb{R}$ with $N \geq 1$ be given. We say that (M,d,m) satisfies the reduced curvature-dimension condition - denoted by $\mathsf{CD}^*(K,N)$ - if and only if for each pair $\nu_0, \nu_1 \in \mathcal{P}_2(M,d,m)$ there exists an optimal coupling π of $\nu_0 = \varrho_0 m$ and $\nu_1 = \varrho_1 m$, and a geodesic $\Gamma : [0,1] \to \mathcal{P}_2(M,d,m)$ connecting ν_0 and ν_1 such that (4.1.3) holds true for all $t \in [0,1]$ and all $N' \geq N$ with the coefficients $\tau_{K,N}^{(t)}(d(x_0,x_1))$ and $\tau_{K,N}^{(1-t)}(d(x_0,x_1))$ replaced by $\sigma_{K,N}^{(t)}(d(x_0,x_1))$ and $\sigma_{K,N}^{(1-t)}(d(x_0,x_1))$, respectively.

The definition of $\mathsf{CD}_{loc}(K,N)$.

Definition 4.5 (Local Curvature-Dimension condition)**.** Let two number $K, N \in \mathbb{R}$ with $N \geq 1$ be given. We say that (M,d,m) satisfies the curvature-dimension condition locally - denoted by $\mathsf{CD}_{loc}(K,N)$ - if and only if each point $x \in M$ has a neighborhood $M(x)$ such that for each pair $\nu_0, \nu_1 \in \mathcal{P}_2(M,d,m)$

supported in $M(x)$ there exists an optimal coupling π of $\nu_0 = \varrho_0 m$ and $\nu_1 = \varrho_1 m$, and a geodesic $\Gamma : [0,1] \to \mathcal{P}_2(M,d,m)$ connecting ν_0 and ν_1 with

$$\begin{aligned}\mathcal{E}_{N'}(\Gamma(t)|m) \le -\int_{M\times M} \Big[&\tau_{K,N'}^{(1-t)}(d(x_0,x_1))\varrho_0^{-1/N'}(x_0) \\ &+ \tau_{K,N'}^{(t)}(d(x_0,x_1))\varrho_1^{-1/N'}(x_1)\Big]\pi(dx_0dx_1),\end{aligned} \tag{4.1.4}$$

for all $t \in [0,1]$ and all $N' \ge N$.

Notice that the geodesic Γ of the above definition can exit from the neighborhood $M(x)$.

If a non-branching metric measure space (M,d,m) satisfies $\mathsf{CD}(K,N)$ then the uniqueness of geodesics can be proven. The next result is taken from [27].

Lemma 4.6. *Assume that (M,d,m) is non-branching an satisfies $\mathsf{CD}(K,N)$ for some pair (K,N). Then for every $x \in \operatorname{supp}[m]$ and m-a.e. $y \in M$ (with the exceptional set depending on x) there exists a unique geodesic between x and y.*

Moreover, there exists a measurable map $\gamma : M^2 \to \mathcal{G}(M)$ such that for $m\otimes m$-a.e. $(x,y) \in M^2$ the curve $t \mapsto \gamma_t(x,y)$ is the unique geodesic connecting x and y.

In the setting of non-branching metric measure space $\mathsf{CD}(K,N)$ has an equivalent point-wise formulation: (M,d,m) satisfies $\mathsf{CD}(K,N)$ if and only if for each pair $\nu_0,\nu_1 \in \mathcal{P}_2(M,d,m)$ and each optimal coupling π of them

$$\varrho_t(\gamma_t(x_0,x_1)) \le \Big[\tau_{K,N'}^{(1-t)}(d(x_0,x_1))\varrho_0^{-1/N'}(x_0) + \tau_{K,N'}^{(t)}(d(x_0,x_1))\varrho_1^{-1/N'}(x_1)\Big]^{-N}, \tag{4.1.5}$$

for all $t \in [0,1]$, and π-a.e. $(x_0,x_1) \in M\times M$. Here ϱ_t is the density of the push-forward of π under the map $(x_0,x_1) \mapsto \gamma_t(x_0,x_1)$.

We recall the definition of the measure contraction property.

A Markov kernel on M is a map $Q : M \times \mathcal{B}(M) \to [0,1]$ (where $\mathcal{B}(M)$ denotes the Borel σ-algebra of M) with the following properties:

(i) for each $x \in M$ the map $Q(x,\cdot) : \mathcal{B}(M) \to [0,1]$ is a probability measure on M;

(ii) for each $A \in \mathcal{B}(M)$ the function $Q(\cdot,A) : M \to [0,1]$ is m-measurable.

Definition 4.7 (Measure contraction property)**.** Let two number $K,N \in \mathbb{R}$ with $N \ge 1$ be given. We say that (M,d,m) satisfies the *measure contraction property* $\mathsf{MCP}(K,N)$ if and only if for each $0 < t < 1$ there exists a Markov kernel Q_t from M^2 to M such that for m^2-a.e. $(x,y) \in M$ and for $Q_t(x,y;\cdot)$-a.e. z the point z is a t-intermediate point of x and y, and such that for m-a.e. $x \in M$ and for every measurable $B \subset M$,

$$\begin{aligned}\int_M \varsigma_{K,N}^{(t)}(d(x,y))Q_t(x,y;B)m(dy) \le m(B),\\ \int_M \varsigma_{K,N}^{(1-t)}(d(x,y))Q_t(y,x;B)m(dy) \le m(B).\end{aligned} \tag{4.1.6}$$

4.2 Polar coordinates

From now on we will assume (M,d,m) to be a non-branching metric measure space satisfying $\mathsf{CD}_{loc}(K,N)$ for some $K,N \in \mathbb{R}$ and $N \ge 1$.

Fix $x_0 \in M$ and a positive $\bar{R}$. Consider the closure of the ball centred in x_0 with radius $\bar{R}$, $\bar{B}_{\bar{R}}(x_0)$ and the family of sets $\{S_p(x_0)\}_{p\le\bar{R}}$. Then the measure $m\llcorner_{\bar{B}_{\bar{R}}(x_0)}$ can be disintegrated in the following way

$$m\llcorner_{\bar{B}_{\bar{R}}(x_0)} = \int \bar{\mathcal{S}}_z q(dz), \qquad q(A) = m(\{x : d(x,x_0) \in A\}).$$

It is fairly easy to prove that the disintegration is strongly consistent. Indeed consider any constant speed geodesic γ going from x_0 to $S_{\bar{R}}(x_0)$ and take $\gamma([0,\bar{R}])$ as the quotient set; clearly $q(\gamma([0,\bar{R}])) = 1$. It follows that the quotient space is a Polish space and then by Corollary 5.4 the disintegration is strongly consistent, i.e.

$$\bar{\mathcal{S}}_{\gamma(p)}(\{x : d(x,x_0) = p\}) = 1, \qquad \text{for } q - a.e.\ \gamma(p) \in \gamma([0,\bar{R}]).$$

Proposition 4.8. *The quotient measure* $q \ll \gamma_\sharp \mathcal{L}^1$.

Proof. Since (M,d,m) satisfies $\mathsf{CD}_{\mathrm{loc}}(K,N)$, it follows from the globalization property of the reduced curvature dimension condition proved in [7], that defining

$$v(r) := m(\bar{B}(x_0)), \qquad s(r) := \limsup_{\delta\to 0} \frac{1}{\delta} m(\bar{B}_{r+\delta}(x_0) \setminus B_r(x_0)),$$

the map $r \mapsto v(r)$ is locally Lipschitz with s as weak derivative, Theorem 2.3. of [27]. Being s the density of $\gamma_\sharp^{-1} q$ w.r.t. $\mathcal{L}^1$, it follows that $\gamma_\sharp^{-1} q \ll \mathcal{L}^1$. $\square$

With a slight abuse of notation $q(dz) = q(z)\gamma_\sharp \mathcal{L}^1$. As direct corollary of Proposition 4.8 we have

$$m\llcorner_{\bar{B}_{\bar{R}}(x_0)} = \int_{\gamma([0,\bar{R}])} \bar{\mathcal{S}}_z q(z)(\gamma_\sharp \mathcal{L}^1)(dz) = \int_{[0,\bar{R}]} \bar{\mathcal{S}}_{\gamma(p)} q(p) \mathcal{L}^1(dp) = \int_{[0,\bar{R}]} \mathcal{S}_p dp.$$

Consider the geodesic $\mathcal{G}_{\bar{R}} : [0,\bar{R}] \to \mathcal{P}_2(M,d)$ going from $\mathcal{G}_{\bar{R}}(0) = \delta_{x_0} \|\mathcal{S}_{\bar{R}}\|$ to $\mathcal{G}_{\bar{R}}(\bar{R}) = \mathcal{S}_{\bar{R}}$.

Lemma 4.9. *The measure* $\mathcal{G}_{\bar{R}}(p)$ *is absolute continuous with respect to the surface measure* $\mathcal{S}_p$.

Proof. Recall that $\mathsf{CD}_{\mathrm{loc}}(K,N)$ implies the Bishop-Gromov volume growth inequality: for all $0 < r \leq R \leq \pi\sqrt{(N-1)/K^*}$

$$\frac{s(r)}{s(R)} \geq \left(\frac{\sin(r\sqrt{K^*/(N-1)})}{\sin(R\sqrt{K^*/(N-1)}} \right)^N, \tag{4.2.1}$$

where $K^* = K(N-1)/N$.

Consider $p \leq \bar{R}$ and A compact set such that $A \subset S_p(x_0)$ and $\mathcal{G}_{\bar{R}}(p)(A) > 0$. Following the proof of Proposition 2.1 and Theorem 2.3 of [27], one can prove (4.2.1) with $s(r)$ replaced by the surface measure of m restricted to $\{\gamma(t) : t \in [0,1], \gamma \in \mathcal{G}(M), \gamma(p/\bar{R}) \in A\}$. Since $\|\mathcal{S}_p\| = s(p)$, the claim follows straightforwardly. $\square$

Hence we have that for $\mathcal{L}^1$-a.e. $p \leq \bar{R}$, $\mathcal{G}_{\bar{R}}(p) = h_{\bar{R}}(\cdot,p)\mathcal{S}_p$.

Remark 4.10. Let us consider the set of geodesic

$$\Gamma_{x_0,S_{\bar{R}}(x_0)} := \{\gamma \in \mathcal{G}(M) : \gamma(0) = x_0, \gamma(1) \in S_{\bar{R}}(x_0)\}.$$

Since (M,d,m) is non-branching, for a fixed $\tau \in (0,1)$ there exists a Borel isomorphism b

$$\{\gamma(\tau) : \gamma \in \Gamma_{x_0,S_{\bar{R}}(x_0)}\} =: Q \ni y \mapsto b(y) := \{\gamma : \gamma(\tau) = y\}.$$

Therefore we can define the following Borel map

$$Q \times [0,\bar{R}] \ni (y,p) \mapsto g(y,p) := b(y)(p/\bar{R}) \in M.$$

Note that g restricted to $Q \times (0,\bar{R})$ is invertible, hence for every $p \in (0,\bar{R})$ consider the following measures:

$$\hat{\mathcal{S}}_p := g_\sharp^{-1} \mathcal{S}_p \in \mathcal{P}(Q \times \{p\}), \qquad g_\sharp^{-1} \mathcal{G}_{\bar{R}}(p) = \hat{h}_{\bar{R}} \hat{\mathcal{S}}_p \in \mathcal{P}(Q \times \{p\}).$$

Consider the measure $\mu := \int \mathcal{G}_{\bar{R}}(p) dp$. It is fairly easy to observe that

$$\mu = g_\sharp \bigg(\int \hat{h}_{\bar{R}} \hat{\mathcal{S}}_p dp \bigg), \qquad \hat{h}_{\bar{R}}(y,p) = h_{\bar{R}}(\gamma(y,p),p).$$

The following regularity result for densities holds true.

Proposition 4.11. *The map* $[0,\bar{R}] \ni p \mapsto \hat{h}_{\bar{R}}(y,p)$ *is locally Lipschitz continuous.*

Proof. *Step 1.* Consider again the measure $\mu := \int \mathcal{G}_{\bar{R}}(p)dp$. Clearly $\mu \ll m$, say $\mu = \rho m$. Let g be the isomorphism introduced in Remark 4.10 and recall that $g(y,\cdot)$ is a geodesic of constant speed equals 1. Define $\hat{\varrho}(y,p) := \varrho(g(y,p))$.

Let us consider the set $\mathcal{T} := \{g(y,p) : (y,p) \in Q \times [0,\bar{R}]\}$. Neglecting a set of m-measure zero, $\mathcal{T}$ is partitioned by the $\{g(y,p) : p \in [0,\bar{R}]\}_{y\in Q}$. Hence we can disintegrate m w.r.t. the aforementioned partition

$$m = g_\sharp\Big(\int B(y,p)\mathcal{L}^1(dp)q_m(dy) \Big), \qquad \int_{[0,\bar{R}]} B(y,p)\mathcal{L}^1(dp) = 1, \ q_m - a.e.\, y \in Q.$$

It is proved in [9] that the map $[0,\bar{R}] \ni p \mapsto B(y,p)$ is strictly positive and locally Lipschitz continuous. Moreover comparing this disintegration with the surface one, we get $g_\sharp\big(B(\cdot,p)q_m\big) = \mathcal{S}_p$.

Step 2. Perform the disintegration of μ w.r.t. the given partition of $\mathcal{T}$:

$$\mu = g_\sharp\Big(\int A(y,p)\mathcal{L}^1(dp)q_\mu(dy) \Big), \qquad \int_{[0,\bar{R}]} A(y,p)\mathcal{L}^1(dp) = 1, \ q_\mu - a.e.\, y \in Q$$

and observe that $\mu(\mathcal{T}) = \|\mathcal{S}_{\bar{R}}\|\bar{R}$. Define the following evolution of sets: for $C \subset \mathcal{T}$ compact set let

$$C_t := \{z \in \mathcal{T} : \exists w \in C, d(x_0,z) = (1-t)d(x_0,w), \exists y \in Q, z,w \in g(y,[0,\bar{R}])\}.$$

The measurability of C_t follows from the measurability of g. Observe that

$$\mu(C_t) = \int \mathcal{G}_R(p)(C_t)dp = \int \mathcal{G}_{\bar{R}}\Big(\frac{p}{1-t}\Big)(C)dp = (1-t)\mu(C).$$

Hence $[0,\bar{R}] \ni p \mapsto A(y,p)$ is constant and therefore $A(y,p) = 1/\bar{R}$ for all t. It follows that

$$\frac{\hat{\rho}(y,p)B(y,p)}{\int \hat{\rho}(y,p)B(y,p)\mathcal{L}^1(dp)} = A(y,p) = \frac{1}{\bar{R}},$$

hence

$$[0,\bar{R}] \ni p \mapsto \hat{h}_{\bar{R}}(y,p) = \hat{\rho}(y,p) = \frac{\int \hat{\rho}(y,p)B(y,p)\mathcal{L}^1(dp)}{B(y,p)\bar{R}}$$

is locally Lipschitz continuous. □

4.3 The $(N-1)$-dimensional estimate

We will derive an $N-1$ version of $\mathsf{CD}_{loc}(K,N)$ for the geodesic $[0,\bar{R}] \ni p \mapsto \mathcal{G}_{\bar{R}}(p)$. For every $r < R$ we define the following map

$$S_r \ni z \mapsto g_{r\to R}(z) := g(P_1(g^{-1}(z)),R) \in S_R.$$

Clearly $g_{r\to R\,\sharp}\mathcal{G}_{\bar{R}}(r) = \mathcal{G}_{\bar{R}}(R)$.

4.3.1 Approximating measures

Fix $0 < r < \bar{R}$ and a positive $t \in (0,1)$. Fix $x \in S_r$ and denote with $M(x)$ the neighborhood from Definition 4.5 in which $\mathsf{CD}_{loc}(K,N)$ can be used. Clearly there exists $\eta > 0$ such that $M(x) \supset B_\eta(x)$. Fix $R < \bar{R}$ such that $0 < R - r \leq \eta$, take $0 < \bar{\eta} < \eta$ and consider the following sets:

$$A := g^{-1}_{r\to R}\Big(g_{r\to R}(B_{\bar{\eta}}(x)) \cap B_{\bar{\eta}}(x)\Big).$$

Possibly restricting to a subset of A, we can assume w.l.o.g. that $1/D \le \hat{h}(y,r), \hat{h}(y,R) \le D$ for every $y \in Q_A := \{y \in Q : g(y,r) \in A\}$. Since we are interested on a local analysis near x we assume that

$$\mathcal{G}_{\bar{R}}(p) = \mathcal{G}_{\bar{R}}(p)\llcorner_{\{g(y,[0,\bar{R}]):y\in Q_A\}}.$$

For every $\varepsilon > 0$ there exits $\delta(\varepsilon,y)$ such that the following equality is satisfied for all $y \in A$

$$\frac{(1-t)\varepsilon}{\hat{h}(y,r)^{-1/(N-1)}\sigma^{(1-t)}_{K,N-1}(R-r)} = \frac{t\delta(\varepsilon,y)}{\hat{h}(y,R)^{-1/(N-1)}\sigma^{(t)}_{K,N-1}(R-r)}. \tag{4.3.1}$$

Throughout the following, whenever h is considered as function of variables in $A \times [0,\bar{p}]$ we intend $h \circ g$.

Now we can define the auxiliary measures to obtain a surface version of $\mathsf{CD}_{loc}(K,N)$. Let

$$\delta(\varepsilon) = \inf\{\delta(\varepsilon,y) : y \in A\}.$$

Possibly restricting to a subset of A, $\delta(\varepsilon) > 0$. Define

$$\begin{aligned}
\mu_{0,\varepsilon} &:= \frac{1}{\varepsilon}\int_{(r-\varepsilon/2,r+\varepsilon/2)} \mathcal{G}_{\bar{R}}(p)dp = \frac{1}{\varepsilon}\int_{(r-\varepsilon/2,r+\varepsilon/2)} \Big(h_{\bar{R}}(\cdot,p)\mathcal{S}_p\Big)dp \\
\mu_{1,\varepsilon} &:= \frac{1}{\delta(\varepsilon)}\int_{(R-\delta(\varepsilon)/2,R+\delta(\varepsilon)/2)} \mathcal{G}_{\bar{R}}(p)dp = \frac{1}{\delta(\varepsilon)}\int_{(R-\delta(\varepsilon)/2,R+\delta(\varepsilon)/2)} \Big(h_{\bar{R}}(\cdot,p)\mathcal{S}_p\Big)dp
\end{aligned} \tag{4.3.2}$$

Observe that for sufficiently small ε both measures are supported in $M(x)$, $\mu_{0,\varepsilon}(B_\eta(x)) = \mu_{1,\varepsilon}(B_\eta(x))$ and $\mu_{0,\varepsilon}, \mu_{1,\varepsilon} \ll m$. Observe that

$$\mu_{0,\varepsilon} = g_\sharp\Big(\frac{1}{\varepsilon}\int_{(r-\varepsilon/2,r+\varepsilon/2)} \hat{h}_{\bar{R}}(\cdot)\hat{\mathcal{S}}_p dp\Big), \qquad \mu_{1,\varepsilon} = g_\sharp\Big(\frac{1}{\delta(\varepsilon)}\int_{(R-\delta(\varepsilon)/2,R+\delta(\varepsilon)/2)} \hat{h}_{\bar{R}}(\cdot)\hat{\mathcal{S}}_p dp\Big). \tag{4.3.3}$$

We assume $K \ge 0$. Denote with $[0,1] \ni t \mapsto \mu_{t,\varepsilon} \in \mathcal{P}_2(M,d,m)$ the geodesic from $\mu_{0,\varepsilon}$ to $\mu_{1,\varepsilon}$ and

$$\mu_{0,\varepsilon} = \varrho_{0,\varepsilon}m, \qquad \mu_{\tau,\varepsilon} = \varrho_{\tau,\varepsilon}m, \qquad \mu_{1,\varepsilon} = \varrho_{1,\varepsilon}m.$$

Let π_ε be the optimal coupling between μ_ε and $\mu_{1,\varepsilon}$ of the form $(Id,T_\varepsilon)_\sharp\mu_\varepsilon$. It follows from $\mathsf{CD}_{loc}(K,N)$, that for π_ε-a.e. $(z_0,z_1) \in M^2$

$$\begin{aligned}
&\varrho^{-1/N}_{t,\varepsilon}(\gamma_t(z_0,z_1)) \\
&\qquad \ge \varrho^{-1/N}_{0,\varepsilon}(z_0)\tau^{(1-t)}_{K,N}(d(z_0,z_1)) + \varrho^{-1/N}_{1,\varepsilon}(x_1)\tau^{(t)}_{K,N}(d(z_0,z_1)).
\end{aligned} \tag{4.3.4}$$

Theorem 4.12. *Fix $(\bar{y},\bar{r})$ and $M(g(\bar{y},\bar{r}))$, then for $\mathcal{S}_r$-a.e $g(y,r) \in M(g(\bar{y},\bar{r}))$ and $g(y,R) \in M(g(\bar{y},\bar{r}))$ the following holds true:*

$$\begin{aligned}
&h_{\bar{R}}^{-\frac{1}{N-1}}(\gamma_t(g(y,r),g(y,R))) \\
&\qquad \ge h_{\bar{R}}(g(y,r))^{-\frac{1}{N-1}}\sigma^{(1-t)}_{K,N-1}(R-r) + h_{\bar{R}}(g(y,R))^{-\frac{1}{N-1}}\sigma^{(t)}_{K,N-1}(R-r).
\end{aligned} \tag{4.3.5}$$

Proof. During the first two steps of the proof we will omit the subscript ε referred to measures and $\bar{R}$ referred to the measure $\mathcal{G}_{\bar{R}}$. Let us expand the right hand side of (4.3.4):

$$\begin{aligned}
&\varrho^{-1/N}_{t,\varepsilon}(\gamma_t(z_0,z_1)) \\
&\qquad \ge \Big(\frac{1}{\varepsilon}h(z_0)\Big)^{-1/N}(1-t)^{1/N}\sigma^{(1-t)}_{K,N-1}(d(z_0,T_\varepsilon(z_0)))^{\frac{N-1}{N}} \\
&\qquad + \Big(\frac{1}{\delta(\varepsilon)}h(T_\varepsilon(z_0))\Big)^{-1/N}t^{1/N}\sigma^{(t)}_{K,N-1}(d(z_0,T_\varepsilon(z_0)))^{\frac{N-1}{N}}.
\end{aligned}$$

Step 1. Let $f:\mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R}$ be defined by the following expression

$$f\left(\begin{pmatrix}a_1\\a_2\end{pmatrix},\begin{pmatrix}b_1\\b_2\end{pmatrix}\right):=a_1b_1+a_2b_2-(a_1^p+a_2^p)^{1/p}(b_1^q+b_2^q)^{1/q},\qquad p=N,q=\frac{N}{N-1}.$$

Recall that $f=0$ if and only if $a_1^p/b_1^q=a_2^p/b_2^q$. Consider

$$\hat a_1:=(\varepsilon(1-t))^{1/N},\qquad \hat b_1:=\hat h(y,r)^{-1/N}\sigma^{(1-t)}_{K,N-1}(R-r)^{\frac{N-1}{N}},$$
$$\hat a_2:=(\delta(\varepsilon,y)t)^{1/N},\qquad \hat b_2:=\hat h(y,R)^{-1/N}\sigma^{(t)}_{K,N-1}(R-r)^{\frac{N-1}{N}}.$$

Then (4.3.1) implies $f(\hat a_i,\hat b_i)=0$. We write the Taylor expansion of f centered at the point $(\hat a_i,\hat b_i)$:

$$\begin{aligned}a_1b_1+a_2b_2=&(a_1^p+a_2^p)^{1/p}(b_1^q+b_2^q)^{1/q}+\langle\nabla f(\hat a_i,\hat b_i),((a_i-\hat a_i),(b_i-\hat b_i)\rangle\\&+O(\|(a_i-\hat a_i,b_i-\hat b_i)\|^2).\end{aligned}$$

and we are interested in considering the above formula at points:

$$a_1:=\hat a_1=(\varepsilon(1-t))^{1/N},\qquad b_1:=h(x_0)^{-1/N}\sigma^{(1-t)}_{K,N-1}(d(x_0,T_\varepsilon(x_0)))^{\frac{N-1}{N}},$$
$$a_2:=(\delta(\varepsilon)t)^{1/N},\qquad b_2:=h(T_\varepsilon(x_0))^{-1/N}\sigma^{(t)}_{K,N-1}(d(x_0,T_\varepsilon(x_0)))^{\frac{N-1}{N}}.$$

Since the derivatives of f have the following expression

$$\frac{\partial f}{\partial a_i}(\hat a_i,\hat b_i)=\hat b_i-\hat a_i^{p-1}\left(\frac{\hat b_1^q+\hat b_2^q}{\hat a_1^p+\hat a_2^p}\right)^{1/q},\qquad \frac{\partial f}{\partial b_i}(\hat a_i,\hat b_i)=\hat a_i-\hat b_i^{q-1}\left(\frac{\hat a_1^p+\hat a_2^p}{\hat b_1^q+\hat b_2^q}\right)^{1/p},$$

we observe that for $j=1,2$ and $j\neq i$:

$$\begin{aligned}\frac{\partial f}{\partial a_i}(\hat a_i,\hat b_i)&=\hat b_i-\hat a_i^{p-1}\left(\frac{\hat b_1^q+\hat b_2^q}{\hat a_1^p+\hat a_2^p}\right)^{1/q}=b_i-(b_1^q+b_2^q)^{1/q}\left(\frac{a_i^p}{a_1^p+a_2^p}\right)^{1/q}\\&=b_i-(b_1^q+b_2^q)^{1/q}\left(1+\frac{a_j^p}{a_i^p}\right)^{-1/q}\\&=b_i-(b_1^q+b_2^q)^{1/q}(b_1^q+b_2^q)^{-1/q}b_i=0,\end{aligned}\tag{4.3.6}$$

and the same calculation holds true for $\frac{\partial f}{\partial b_i}(\hat a_i,\hat b_i)$. It follows that the Taylor expansion of f centered at $(\hat a_i,\hat b_i)$ and evaluated at point (a_i,b_i) has the following expression:

$$\begin{aligned}&\left(\frac{h(z_0)}{\varepsilon}\right)^{-\frac{1}{N}}(1-t)^{\frac1N}\sigma^{(1-t)}_{K,N-1}(d(z_0,T_\varepsilon(z_0)))^{\frac{N-1}{N}}+\left(\frac{h(T_\varepsilon(z_0))}{\delta(\varepsilon)}\right)^{-\frac1N}t^{\frac1N}\sigma^{(t)}_{K,N-1}(d(z_0,T_\varepsilon(z_0)))^{\frac{N-1}{N}}\\&\quad=\Big(\varepsilon(1-t)+\delta(\varepsilon)t\Big)^{\frac1N}\\&\quad\times\Big(h(z_0)^{-\frac{1}{N-1}}\sigma^{(1-t)}_{K,N-1}(d(z_0,T_\varepsilon(z_0)))+h(T_\varepsilon(z_0))^{-\frac{1}{N-1}}\sigma^{(t)}_{K,N-1}(d(z_0,T_\varepsilon(z_0))\Big)^{\frac{N-1}{N}}\\&\quad+O(\|(a_i-\hat a_i,b_i-\hat b_i)\|^2).\end{aligned}\tag{4.3.7}$$

Step 2. Now we can rewrite (4.3.4). For π_ε a.e. (z_0,z_1) it holds

$$\begin{aligned}&\varrho_{t,\varepsilon}(\gamma_t(z_0,z_1))^{-\frac1N}\\&\quad\geq\Big(\varepsilon(1-t)+\delta(\varepsilon)t\Big)^{\frac1N}\Big(h(z_0)^{-\frac{1}{N-1}}\sigma^{(1-t)}_{K,N-1}(d(z_0,z_1))+h(z_1)^{-\frac{1}{N-1}}\sigma^{(t)}_{K,N-1}(d(z_0,z_1))\Big)^{\frac{N-1}{N}}\\&\quad+O(\|(a_i-\hat a_i,b_i-\hat b_i)\|^2).\end{aligned}$$

It follows that for π_ε-a.e. (z_0,z_1) it holds

$$
\begin{aligned}
\Big((\varepsilon(1-t)+\delta(\varepsilon)t)\varrho_{t,\varepsilon}(\gamma_t(x_0,x_1))\Big)^{-\frac{1}{N-1}}&\\
\geq \Big\{\Big(h(z_0)^{-\frac{1}{N-1}}\sigma^{(1-t)}_{K,N-1}(d(z_0,z_1))&+h(z_1)^{-\frac{1}{N-1}}\sigma^{(t)}_{K,N-1}(d(z_0,z_1))\Big)^{\frac{N-1}{N}}\\
&+O(\|(a_i-\hat a_i,b_i-\hat b_i)\|^2)\Big(\varepsilon(1-t)+\delta(\varepsilon)t\Big)^{-\frac{1}{N}}\Big\}^{\frac{N}{N-1}}.
\end{aligned}
$$

Observe that $\mu_{t,\varepsilon}=\mathcal{L}^1(I_{t,\varepsilon})^{-1}\int_{I_{t,\varepsilon}}\mathcal{G}_p dp$ with $I_{t,\varepsilon}$ convex combination of $(r-\varepsilon/2,r+\varepsilon/2)$ and $(R-\delta(\varepsilon),R+\delta(\varepsilon))$. Hence $\mathcal{L}^1(I_{t,\varepsilon})=\varepsilon(1-t)+\delta(\varepsilon)t$, therefore

$$
(\varepsilon(1-t)+\delta(\varepsilon)t)\varrho_{t,\varepsilon}(\gamma_t(x_0,x_1))=h(\gamma_t(x_0,x_1)).
$$

Let ν be the quotient measure of π_ε w.r.t. the family of disjoint sets $\{g(y,[0,\bar R])\times M\}_{y\in Q}$. Note that ν doesn't depend on ε. It follows that $\pi_\varepsilon=\int\pi_{\varepsilon,y}\nu(dy)$ with $\pi_{\varepsilon,y}\rightharpoonup\delta_{g(y,r),g(y,R)}$. It follows that

$$
\begin{aligned}
\int h(\gamma_t(z_0,z_1))^{-\frac{1}{N-1}}&\pi_{\varepsilon,y}(dz_0dz_1)\nu(dy)\\
\geq \int\Big\{\Big(h(z_0)^{-\frac{1}{N-1}}&\sigma^{(1-t)}_{K,N-1}(d(z_0,z_1))+h(z_1)^{-\frac{1}{N-1}}\sigma^{(t)}_{K,N-1}(d(z_0,z_1))\Big)^{\frac{N-1}{N}}\\
&+O(\|(a_i-\hat a_i,b_i-\hat b_i)\|^2)\Big(\varepsilon(1-t)+\delta(\varepsilon)t\Big)^{-\frac{1}{N}}\Big\}^{\frac{N}{N-1}}\pi_{\varepsilon,y}(dz_0dz_1)\nu(dy).
\end{aligned}
$$

From (4.3.1) and the boundedness of h, $D\leq t\delta(\varepsilon,y)/(1-t)\varepsilon\leq 1/D$ uniformly in ε. Taking advantage of of the Lipschitz regularity of $t\mapsto h(g(y,t))$, letting $\varepsilon\to0$ one can prove that $O(\|(a_i-\hat a_i,b_i-\hat b_i)\|^2)(\varepsilon(1-t)+\delta(\varepsilon)t)^{-\frac{1}{N}}\sim\varepsilon^{1/N}$. Being the only depending on ε, and being the other terms Lipschitz continuous when restricted on geodesics, it follows that

$$
\begin{aligned}
\int h(\gamma_t(z_0,z_1))^{-\frac{1}{N-1}}&\pi(dz_0dz_1)\\
\geq \int\Big\{\Big(h(z_0)^{-\frac{1}{N-1}}&\sigma^{(1-t)}_{K,N-1}(d(z_0,z_1))+h(z_1)^{-\frac{1}{N-1}}\sigma^{(t)}_{K,N-1}(d(z_0,z_1))\Big)^{\frac{N-1}{N}}\Big\}^{\frac{N}{N-1}}\pi(dz_0dz_1).
\end{aligned}
$$

where $\pi_\varepsilon\rightharpoonup\pi\in\Pi(\mu,\nu)$ with $\mu_\varepsilon\rightharpoonup\mu$ and $\nu_\varepsilon\rightharpoonup\nu$. Since A can be taken as small as we want, we obtain the point-wise version of the above inequality and the claim follows. □

4.4 The global estimates

From Theorem 4.12 we have that fixed $y\in Q$: for every $0<r<\bar R$ there exists a neighborhood $M(g(y,r))$ of $g(y,r)$ such that for all $r<R<\bar R$ with $g(y,R)\in M(g(y,r))$ it holds

$$
\begin{aligned}
h_{\bar R}(\gamma_t(g(y,r),g(y,R)))^{-\frac{1}{N-1}}\geq{}& h_{\bar R}(g(y,r))^{-\frac{1}{N-1}}\sigma^{(1-t)}_{K,N-1}(R-r)\\
&+h_{\bar R}(g(y,R))^{-\frac{1}{N-1}}\sigma^{(t)}_{K,N-1}(R-r),
\end{aligned}
$$

for every $t\in[0,1]$. Since the comparison is always on the same geodesic, the map $t\mapsto h_{\bar R}(g(y,t))$ is a real variable map and the above inequality is equivalent to: for every $0<r<\bar R$ there exists ε_r such that if $R-r<\varepsilon_r$, then the following holds true

$$
h_{\bar R}(r+t(R-r)))^{-\frac{1}{N-1}}\geq h_{\bar R}(r)^{-\frac{1}{N-1}}\sigma^{(1-t)}_{K,N-1}(R-r)+h_{\bar R}(R)^{-\frac{1}{N-1}}\sigma^{(t)}_{K,N-1}(R-r), \tag{4.4.1}
$$

for every $t\in[0,1]$.

Lemma 4.13 (Midpoints)**.** *For all* $0 \le r \le R \le \bar{R}$ *inequality* (4.4.1) *holds true for all* $t \in [0,1]$ *if and only if for all* $0 \le r \le R \le \bar{R}$ *we have:*

$$h_{\bar{R}}(\frac{r+R}{2})^{-\frac{1}{N-1}} \ge h_{\bar{R}}(r)^{-\frac{1}{N-1}}\sigma^{(1/2)}_{K,N-1}(R-r) + h_{\bar{R}}(R)^{-\frac{1}{N-1}}\sigma^{(1/2)}_{K,N-1}(R-r). \tag{4.4.2}$$

Proof. We only consider the case $K>0$. The general case requires analogous calculations. Fix $0 \le r \le R \le \bar{R}$ and put $\theta := R-r$.

Step 1. For every $k \in \mathbb{N}$ we have

$$\begin{aligned} h_{\bar{R}}(r+l2^{-k}\theta))^{-\frac{1}{N-1}} \ge\ & h_{\bar{R}}(r+(l-1)2^{-k}\theta))^{-\frac{1}{N-1}}\sigma^{(1/2)}_{K,N-1}(2^{-k+1}\theta) \\ &+ h_{\bar{R}}(r+(l+1)2^{-k}\theta))^{-\frac{1}{N-1}}\sigma^{(1/2)}_{K,N-1}(2^{-k+1}\theta), \end{aligned}$$

for every odd $l = 0, \dots, 2^k$.

Step 2. We perform an induction argument on k: suppose that inequality (4.4.1) is satisfied for all $t = l2^{-k+1} \in [0,1]$ with l odd, then (4.4.1) is verified by every $t = l2^{-k} \in [0,1]$ with l odd:

$$\begin{aligned} &h_{\bar{R}}(r+l2^{-k}\theta))^{-\frac{1}{N-1}} \\ &\quad \ge h_{\bar{R}}(r+(l-1)2^{-k}\theta))^{-\frac{1}{N-1}}\sigma^{(1/2)}_{K,N-1}(2^{-k+1}\theta) \\ &\qquad + h_{\bar{R}}(r+(l+1)2^{-k}\theta))^{-\frac{1}{N-1}}\sigma^{(1/2)}_{K,N-1}(2^{-k+1}\theta) \\ &\quad \ge \sigma^{(1/2)}_{K,N-1}(2^{-k+1}\theta)\Big[h_{\bar{R}}(r)^{-\frac{1}{N-1}}\sigma^{(1-(l-1)2^{-k})}_{K,N-1}(\theta) + h_p(R)^{-\frac{1}{N-1}}\sigma^{((l-1)2^{-k})}_{K,N-1}(\theta)\Big] \\ &\qquad + \sigma^{(1/2)}_{K,N-1}(2^{-k+1}\theta)\Big[h_{\bar{R}}(r)^{-\frac{1}{N-1}}\sigma^{(1-(l+1)2^{-k})}_{K,N-1}(\theta) + h_p(R)^{-\frac{1}{N-1}}\sigma^{((l+1)2^{-k})}_{K,N-1}(\theta)\Big]. \end{aligned}$$

Following the calculation of the proof of Proposition 2.10 of [7], one obtain that

$$h_{\bar{R}}(r+l2^{-k}\theta))^{-\frac{1}{N-1}} \ge h_{\bar{R}}(r)^{-\frac{1}{N-1}}\sigma^{(1-l2^{-k})}_{K,N-1}(\theta) + h_{\bar{R}}(R)^{-\frac{1}{N-1}}\sigma^{(l2^{-k})}_{K,N-1}(\theta).$$

The claim is easily proved by the continuity of $h_{\bar{R}}$ and of σ. □

Theorem 4.14 (Local to Global)**.** *Suppose that for every* $r \in (0,\bar{R})$ *there exists* $\varepsilon_r > 0$ *such that whenever* $R \in B_{\varepsilon(r)}(r)$ *then* (4.4.1) *holds true for any* $t \in [0,1]$*. Then* (4.4.1) *holds true for any* $0 \le r < R < \bar{R}$ *and* $t \in [0,1]$*.*

Proof. We only consider the case $K>0$. The general case requires analogous calculations. Fix $0 < r < R \le \bar{R}$ and $\theta := R-r$.

Step 1. According to our assumption, every point $r \in [0,\bar{R}]$ has a neighborhood $B_{\varepsilon(r)}(r)$ such that if $R \in B_{\varepsilon(r)}(r)$ then (4.4.1) is verified. By compactness of $[0,\bar{R}]$ there exist $x_1,\dots,x_n$ such that the family $\{B_{\varepsilon(x_i)/2}(x_i)\}_{i=1,\dots,n}$ is a covering of $[0,\bar{R}]$. Let $\lambda := \min\{\varepsilon(x_i)/2 : i = 1,\dots,n\}$. Possibly taking a lower value for λ, we assume that $\lambda = 2^{-k}\theta$. Hence we have

$$\begin{aligned} h_{\bar{R}}(r+\frac{1}{2}\theta)^{-\frac{1}{N-1}} \ge\ & h_{\bar{R}}(r+\frac{1}{2}\theta-2^{-k}\theta)^{-\frac{1}{N-1}}\sigma^{(1/2)}_{K,N-1}(2^{-k+1}\theta) \\ &+ h_{\bar{R}}(r+\frac{1}{2}\theta+2^{-k}\theta)^{-\frac{1}{N-1}}\sigma^{(1/2)}_{K,N-1}(2^{-k+1}\theta). \end{aligned}$$

Step 2. We iterate the above inequality:

$$
\begin{aligned}
h_{\bar{R}}(r+\frac{1}{2}\theta)^{-\frac{1}{N-1}} \geq\ & h_{\bar{R}}(r+\frac{1}{2}\theta-2^{-k}\theta)^{-\frac{1}{N-1}}\sigma^{(1/2)}_{K,N-1}(2^{-k+1}\theta) \\
& + h_{\bar{R}}(r+\frac{1}{2}\theta+2^{-k}\theta)^{-\frac{1}{N-1}}\sigma^{(1/2)}_{K,N-1}(2^{-k+1}\theta) \\
\geq\ & \sigma^{(1/2)}_{K,N-1}(2^{-k+1}\theta)\Big[h_{\bar{R}}(r+\frac{1}{2}\theta-2^{-k+1}\theta)^{-\frac{1}{N-1}}\sigma^{(1/2)}_{K,N-1}(2^{-k+1}\theta) \\
& + h_{\bar{R}}(r+\frac{1}{2}\theta)^{-\frac{1}{N-1}}\sigma^{(1/2)}_{K,N-1}(2^{-k+1}\theta)\Big] \\
& + \sigma^{(1/2)}_{K,N-1}(2^{-k+1}\theta)\Big[h_{\bar{R}}(r+\frac{1}{2}\theta)^{-\frac{1}{N-1}}\sigma^{(1/2)}_{K,N-1}(2^{-k+1}\theta) \\
& + h_{\bar{R}}(r+\frac{1}{2}\theta+2^{-k+1}\theta)^{-\frac{1}{N-1}}\sigma^{(1/2)}_{K,N-1}(2^{-k+1}\theta)\Big] \\
\geq\ & \sigma^{(1/2)}_{K,N-1}(2^{-k+1}\theta)^2 h_{\bar{R}}(r+\frac{1}{2}\theta-2^{-k+1}\theta)^{-\frac{1}{N-1}} \\
& + \sigma^{(1/2)}_{K,N-1}(2^{-k+1}\theta)^2 h_{\bar{R}}(r+\frac{1}{2}\theta+2^{-k+1}\theta)^{-\frac{1}{N-1}}.
\end{aligned}
$$

Observing that $\sigma^{(1/2)}_{K,N-1}(\alpha)^2 \geq \sigma^{(1/2)}_{K,N-1}(2\alpha)$, it is fairly easy to obtain:

$$
\begin{aligned}
h_{\bar{R}}(r+\frac{1}{2}\theta)^{-\frac{1}{N-1}} \geq\ & h_{\bar{R}}(r+\frac{1}{2}\theta-2^{-k+i}\theta)^{-\frac{1}{N-1}}\sigma^{(1/2)}_{K,N-1}(2^{-k+i+1}\theta) \\
& + h_{\bar{R}}(r+\frac{1}{2}\theta+2^{-k+i}\theta)^{-\frac{1}{N-1}}\sigma^{(1/2)}_{K,N-1}(2^{-k+i+1}\theta),
\end{aligned}
$$

for every $i = 0, \dots, k$. For $i = k$ Lemma 4.13 implies the claim. $\square$

4.5 From local CD(K,N) to global MCP(K,N)

So we have proved that the density $h_{\bar{R}}(\cdot)$ of the geodesic $[0,\bar{R}] \ni p \mapsto \mathcal{G}_{\bar{R}}(p)$ satisfy the following inequality:

$$
\begin{aligned}
h_{\bar{R}}(\gamma_t(g(y,r),g(y,R)))^{-\frac{1}{N-1}} \geq\ & h_{\bar{R}}(g(y,r))^{-\frac{1}{N-1}}\sigma^{(1-t)}_{K,N-1}(R-r) \\
& + h_{\bar{R}}(g(y,R))^{-\frac{1}{N-1}}\sigma^{(t)}_{K,N-1}(R-r).
\end{aligned}
$$

for every $y \in Q$ and $0 < r < R \leq \bar{R}$.

Fix a generic $x_0 \in M$ and consider the following measure

$$
\mu = \int_{[\bar{r},\bar{R}]} \mathcal{S}_p dp.
$$

Let $[0,1] \ni t \mapsto \mu_t \in \mathcal{P}_2(M,d,m)$ be the geodesic such that $\mu_0 = \mu$ and $\mu_1 = \delta_{x_0}\|\mu\|$ with $\mu_t = \varrho_t m$. Let moreover $\pi_t \in \Pi(\mu_0,\mu_t)$ the corresponding optimal coupling.

Proposition 4.15. *Fix $t \in [0,1)$. Then for π_t-a.e. $(z_0,z_1) \in M^2$ the following holds true*

$$
\varrho_{ts}(\gamma_s(z_0,z_1))^{-1/N} \geq \varrho_0(z_0)^{-1/N}\tau^{(1-s)}_{K,N}(d(z_0,z_1)) + \varrho_t(z_1)^{-1/N}\tau^{(s)}_{K,N}(d(z_0,z_1)),
$$

for every $s \in [0,1]$.

Proof. Let $[0,1] \ni s \mapsto \Gamma(s)$ geodesic with $\Gamma_0 = \mu, \Gamma(1) = \mu_t$. Since

$$
\mu = \int_{[\bar{r},\bar{R}]} \mathcal{S}_p dp, \qquad \mu_t = \int_{[\bar{r},\bar{R}]} \mathcal{G}_p((1-t)p)dp = \frac{1}{1-t}\int_{(1-t)[\bar{r},\bar{R}]} h_{p/(1-t)}(\cdot,p)\mathcal{S}_p dp,
$$

therefore

$$\Gamma(s) = \mu_{ts} = \frac{1}{1-ts}\int_{(1-ts)[\bar r,\bar R]} h_{p/(1-ts)}(\cdot,p)\mathcal{S}_p dp. \tag{4.5.1}$$

Let $\pi_t \in \Pi(\mu,\mu_t)$ be the optimal plan and consider $\bar r \le \bar p \le \bar R$. and $z_0 \in S_{\bar p}$. Then the unique z_1 such that (z_0,z_1) is in the support of the optimal plan π_t, belongs to $\mathcal{S}_{(1-t)\bar p}$. Then from Theorem 4.12 and (4.5.1)

$$\begin{aligned}
\varrho_{ts}(\gamma_s(x_0,x_1))^{-1/N} &= \Big(\frac{1}{1-ts}h_{\bar p}(\gamma_s(x_0,x_1))\Big)^{-1/N}\\
&= \Big(\frac{1}{(1-t)s+1-s}\Big)^{-\frac{1}{N}}\Big(h_{\bar p}(\gamma_s(x_0,x_1))^{-\frac{1}{N-1}}\Big)^{\frac{N-1}{N}}\\
&\ge (1-s)^{1/N}\Big(h_{\bar p}(x_0)^{-\frac{1}{N-1}}\sigma^{(1-s)}_{K,N-1}(d(x_0,x_1))\Big)^{\frac{N-1}{N}}\\
&+ ((1-t)s)^{1/N}\Big(h_{\bar p}(x_1)^{-\frac{1}{N-1}}\sigma^{(s)}_{K,N-1}(d(x_0,x_1))\Big)^{\frac{N-1}{N}}\\
&= h_{\bar p}(x_0)^{-\frac{1}{N}}\tau^{(1-s)}_{K,N}(d(x_0,x_1)) + \Big(\frac{h_{\bar p}(x_1)}{1-t}\Big)^{-\frac{1}{N}}\tau^{(s)}_{K,N}(d(x_0,x_1))\\
&= \varrho_0(z_0)^{-1/N}\tau^{(1-s)}_{K,N}(d(z_0,z_1)) + \varrho_t(z_1)^{-1/N}\tau^{(s)}_{K,N}(d(z_0,z_1)),
\end{aligned}$$

The claim follows. ☐

So far we have proven that given $\mu_0 := m(A)^{-1}m\llcorner_A$, $x_0 \in \mathrm{supp}[m]$ and the unique geodesic $[0,1]\ni t \mapsto \Gamma(t)$ such that $\Gamma(0)=\mu_0$, $\Gamma(1)=\delta_{x_0}$ and $\Gamma(t)=\varrho_t m$ for $t\in[0,1)$ we have for any $t\in[0,1)$:

$$\begin{aligned}
\mathcal{E}_{N'}(\Gamma(ts)|m) \le -\int_{M\times M}\Big[&\tau^{(1-s)}_{K,N'}(d(x_0,x_1))\varrho_0^{-1/N'}(x_0)\\
&+\tau^{(s)}_{K,N'}(d(x_0,x_1))\varrho_t^{-1/N'}(x_1)\Big]\pi_t(dx_0dx_1),
\end{aligned} \tag{4.5.2}$$

for all $s\in[0,1]$ and all $N'\ge N$, where $\pi_t = (P_0,P_t)_\sharp\Xi$.

We are ready to prove the main theorem of this chapter.

Theorem 4.16. *Let (M,d,m) be a non-branching metric measure spaces satisfying* $\mathsf{CD}_{loc}(K,N)$. *Then (M,d,m) satisfies* $\mathsf{MCP}(K,N)$.

Proof. Step 1. Let $\gamma : M^2 \to \mathcal{G}(M)$ be the map introduced in Lemma 4.6 and define for each $t\in[0,1]$ a Markov kernel Q_t from M^2 to M by

$$Q_t(x,y;B) := 1_B(\gamma_t(x,y))$$

and for each pair t,x a measure $m_{t,x} = \int Q_t(x,y;\cdot)m(dy)$.

For each $x\in M$ let M_x denote the ser of all $y\in M$ for which there exists a unique geodesic connecting x and y and let M_0 be the set of x such that $m(M\setminus M_x)=0$. By assumption $m(M\setminus M_0)=0$.

Step 2. Fix $x_0\in M_0$ and $B\subset M$. Put $A_0 := \gamma_t(x_0,\cdot)^{-1}(B)$ and $\mu_0 := m(A_0)^{-1}m\llcorner_{A_0}$. Considering $s=1$ in (4.5.2) it follows that

$$m(B)^{1/N} \ge \inf_{y\in A_0}\tau^{(t)}_{K,N}(d(y,x_0))m(A_0)^{1/N},$$

or equivalently

$$m(B) \ge \inf_{y\in\gamma_t(x_0,\cdot)^{-1}(B)}\varsigma^{(t)}_{K,N}(d(y,x_0))m(\gamma_t(x_0,\cdot)^{-1}(B)) = \inf_{z\in B}\varsigma^{(t)}_{K,N}\Big(\frac{d(z,x_0)}{t}\Big)m_{t,x_0}(B).$$

Decomposing B into a disjoint union $\cup_i B_i$ with $B_i = B \cap (\bar{B}_{\varepsilon i}(x_0) \setminus \bar{B}_{\varepsilon(i-1)}(x_0)$, and applying the previous estimate to each of the B_i we obtain as $\varepsilon \to 0$

$$m(B) \geq \int_B \varsigma_{K,N}^{(t)} \left(\frac{d(z,x_0)}{t} \right) m_{t,x_0}(dz)$$

or equivalently

$$m(B) \geq \int_B \varsigma_{K,N}^{(t)} (d(z,x_0)) Q_t(x_0,y;B) m(dy).$$

□

Chapter 5

Appendix

In this section we recall some general facts about projective classes, the Disintegration Theorem for measures, measurable selection principles, geodesic spaces and optimal transportation problems.

5.1 Borel, projective and universally measurable sets

The *projective class* $\Sigma^1_1(X)$ is the family of subsets A of the Polish space X for which there exists Y Polish and $B \in \mathcal{B}(X \times Y)$ such that $A = P_1(B)$. The *coprojective class* $\Pi^1_1(X)$ is the complement in X of the class $\Sigma^1_1(X)$. The class Σ^1_1 is called *the class of analytic sets*, and Π^1_1 are the *coanalytic sets*.

The *projective class* $\Sigma^1_{n+1}(X)$ is the family of subsets A of the Polish space X for which there exists Y Polish and $B \in \Pi^1_n(X \times Y)$ such that $A = P_1(B)$. The *coprojective class* $\Pi^1_{n+1}(X)$ is the complement in X of the class Σ^1_{n+1}.

If Σ^1_n, Π^1_n are the projective, coprojective pointclasses, then the following holds (Chapter 4 of [25]):

1. Σ^1_n, Π^1_n are closed under countable unions, intersections (in particular they are monotone classes);

2. Σ^1_n is closed w.r.t. projections, Π^1_n is closed w.r.t. coprojections;

3. if $A \in \Sigma^1_n$, then $X \setminus A \in \Pi^1_n$;

4. the *ambiguous class* $\Delta^1_n = \Sigma^1_n \cap \Pi^1_n$ is a σ-algebra and $\Sigma^1_n \cup \Pi^1_n \subset \Delta^1_{n+1}$.

We will denote by $\mathcal{A}$ the σ-algebra generated by Σ^1_1: clearly $\mathcal{B} = \Delta^1_1 \subset \mathcal{A} \subset \Delta^1_2$.

We recall that a subset of X Polish is *universally measurable* if it belongs to all completed σ-algebras of all Borel measures on X: it can be proved that every set in $\mathcal{A}$ is universally measurable. We say that $f : X \to \mathbb{R} \cup \{\pm\infty\}$ is a *Souslin function* if $f^{-1}(t, +\infty] \in \Sigma^1_1$.

Lemma 5.1. *If $f : X \to Y$ is universally measurable, then $f^{-1}(U)$ is universally measurable if U is.*

Proof. If $\mu \in \mathcal{M}(X)$, then $f_\sharp\mu \in \mathcal{M}(Y)$, so for $U \subset Y$ universally measurable there exist Borel sets B, B' such that $B \subset U \subset B'$ and

$$0 = (f_\sharp\mu)(B' \setminus B) = \mu\big(f^{-1}(B') \setminus f^{-1}(B)\big).$$

Since $f^{-1}(B), f^{-1}(B') \subset X$ are universally measurable, there exists Borel sets C, C' such that

$$C \subset f^{-1}(B) \subset f^{-1}(U) \subset f^{-1}(B') \subset C'$$

and $\mu(C' \setminus C) = 0$. The conclusion follows. □

5.2 Disintegration of measures

Given a measurable space $(R, \mathscr{R})$ and a function $r : R \to S$, with S generic set, we can endow S with the *push forward σ-algebra $\mathscr{S}$* of $\mathscr{R}$:

$$Q \in \mathscr{S} \quad \Longleftrightarrow \quad r^{-1}(Q) \in \mathscr{R},$$

which could be also defined as the biggest σ-algebra on S such that r is measurable. Moreover given a measure space $(R, \mathscr{R}, \rho)$, the *push forward measure* η is then defined as $\eta := (r_\sharp \rho)$.

Consider a probability space $(R, \mathscr{R}, \rho)$ and its push forward measure space $(S, \mathscr{S}, \eta)$ induced by a map r. From the above definition the map r is clearly measurable.

Definition 5.2. A *disintegration* of ρ *consistent with* r is a map $\rho : \mathscr{R} \times S \to [0, 1]$ such that

1. $\rho_s(\cdot)$ is a probability measure on $(R, \mathscr{R})$ for all $s \in S$,

2. $\rho_\cdot(B)$ is η-measurable for all $B \in \mathscr{R}$,

and satisfies for all $B \in \mathscr{R}, C \in \mathscr{S}$ the consistency condition

$$\rho\left(B \cap r^{-1}(C)\right) = \int_C \rho_s(B)\eta(ds).$$

A disintegration is *strongly consistent with respect to* r if for all s we have $\rho_s(r^{-1}(s)) = 1$.

The measures ρ_s are called *conditional probabilities*.

We say that a σ-algebra $\mathcal{H}$ is *essentially countably generated* with respect to a measure m if there exists a countably generated σ-algebra $\hat{\mathcal{H}}$ such that for all $A \in \mathcal{H}$ there exists $\hat{A} \in \hat{\mathcal{H}}$ such that $m(A \triangle \hat{A}) = 0$.

We recall the following version of the disintegration theorem that can be found on [21], Section 452 (see [8] for a direct proof).

Theorem 5.3 (Disintegration of measures)**.** *Assume that $(R, \mathscr{R}, \rho)$ is a countably generated probability space, $R = \{R_s\}_{s \in S}$ a partition of R, $r : R \to S$ the quotient map and $(S, \mathscr{S}, \eta)$ the quotient measure space. Then $\mathscr{S}$ is essentially countably generated w.r.t. η and there exists a unique disintegration $s \mapsto \rho_s$ in the following sense: if ρ_1, ρ_2 are two consistent disintegration then $\rho_{1,s}(\cdot) = \rho_{2,s}(\cdot)$ for η-a.e. s.*

If $\{S_n\}_{n \in \mathbb{N}}$ is a family essentially generating $\mathscr{S}$ define the equivalence relation:

$$s \sim s' \iff \{s \in S_n \iff s' \in S_n, \ \forall\, n \in \mathbb{N}\}.$$

Denoting with p the quotient map associated to the above equivalence relation and with $(L, \mathscr{L}, \lambda)$ the quotient measure space, the following properties hold:

- *$\hat{R}_l := \cup_{s \in p^{-1}(l)} R_s = (p \circ r)^{-1}(l)$ is ρ-measurable and $R = \cup_{l \in L} \hat{R}_l$;*

- *the disintegration $\rho = \int_L \rho_l \lambda(dl)$ satisfies $\rho_l(\hat{R}_l) = 1$, for λ-a.e. l. In particular there exists a strongly consistent disintegration w.r.t. $p \circ r$;*

- *the disintegration $\rho = \int_S \rho_s \eta(ds)$ satisfies $\rho_s = \rho_{p(s)}$ for η-a.e. s.*

In particular we will use the following corollary.

Corollary 5.4. *If $(S, \mathscr{S}) = (X, \mathcal{B}(X))$ with X Polish space, then the disintegration is strongly consistent.*

5.3 Selection principles

Given a multivalued function $F : X \to Y$, X, Y metric spaces, the *graph* of F is the set

$$\mathrm{graph}(F) := \big\{(x,y) : y \in F(x)\big\}. \tag{5.3.1}$$

The *inverse image* of a set $S \subset Y$ is defined as:

$$F^{-1}(S) := \big\{x \in X \ : \ F(x) \cap S \neq \emptyset\big\}. \tag{5.3.2}$$

For $F \subset X \times Y$, we denote also the sets

$$F_x := F \cap \{x\} \times Y, \quad F^y := F \cap X \times \{y\}. \tag{5.3.3}$$

In particular, $F(x) = P_2(\mathrm{graph}(F)_x)$, $F^{-1}(y) = P_1(\mathrm{graph}(F)^y)$. We denote by F^{-1} the graph of the inverse function

$$F^{-1} := \big\{(x,y) : (y,x) \in F\big\}. \tag{5.3.4}$$

We say that F is *$\mathcal{R}$-measurable* if $F^{-1}(B) \in \mathcal{R}$ for all B open. We say that F is *strongly Borel measurable* if inverse images of closed sets are Borel. A multivalued function is called *upper-semicontinuous* if the preimage of every closed set is closed: in particular u.s.c. maps are strongly Borel measurable.

In the following we will not distinguish between a multifunction and its graph. Note that the *domain of F* (i.e. the set $P_1(F)$) is in general a subset of X. The same convention will be used for functions, in the sense that their domain may be a subset of X.

Given $F \subset X \times Y$, a *section u of F* is a function from $P_1(F)$ to Y such that $\mathrm{graph}(u) \subset F$. We recall the following selection principle, Theorem 5.5.2 of [25], page 198.

Theorem 5.5. *Let X and Y be Polish spaces, $F \subset X \times Y$ analytic, and $\mathcal{A}$ the σ-algebra generated by the analytic subsets of X. Then there is an $\mathcal{A}$-measurable section $u : P_1(F) \to Y$ of F.*

A *cross-section of the equivalence relation E* is a set $S \subset E$ such that the intersection of S with each equivalence class is a singleton. We recall that a set $A \subset X$ is saturated for the equivalence relation $E \subset X \times X$ if $A = \cup_{x\in A} E(x)$.

The next result is taken from [25], Theorem 5.2.1.

Theorem 5.6. *Let Y be a Polish space, X a nonempty set, and $\mathcal{L}$ a σ-algebra of subset of X. Every $\mathcal{L}$-measurable, closed value multifunction $F : X \to Y$ admits an $\mathcal{L}$-measurable section.*

We will use the following corollary.

Corollary 5.7. *Let $F \subset X \times X$ be $\mathcal{A}$-measurable, X Polish, such that F_x is closed for every $x \in X$ and define the equivalence relation $x \sim y \ \Leftrightarrow \ F(x) = F(y)$. Then there exists a $\mathcal{A}$-section $f : P_1(F) \to X$ such that $(x, f(x)) \in F$ and $f(x) = f(y)$ if $x \sim y$.*

Proof. For all open sets $G \subset X$, consider the sets $F^{-1}(G) = P_1(F \cap X \times G) \in \mathcal{A}$, and let $\mathcal{R}$ be the σ-algebra generated by $F^{-1}(G)$. Clearly $\mathcal{R} \subset \mathcal{A}$.

If $x \sim y$, then

$$x \in F^{-1}(G) \quad \Longleftrightarrow \quad y \in F^{-1}(G),$$

so that each equivalence class is contained in an atom of $\mathcal{R}$, and moreover by construction $x \mapsto F(x)$ is $\mathcal{R}$-measurable.

We thus conclude by using Theorem 5.6 that there exists an $\mathcal{R}$-measurable section f: this measurability condition implies that f is constant on atoms, in particular on equivalence classes. □

5.4 General facts about optimal transportation

Let $(X, \mathcal{B}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be two Polish probability spaces and $c : X \times Y \to \mathbb{R}$ be a Borel measurable function. Consider the set of *transference plans*

$$\Pi(\mu, \nu) := \Big\{\pi \in \mathcal{P}(X \times Y) : (P_1)_\sharp \pi = \mu, (P_2)_\sharp \pi = \nu\Big\}.$$

Define the functional

$$\begin{array}{rcl} \mathcal{I} : \Pi(\mu,\nu) & \to & \mathbb{R}^+ \\ \pi & \mapsto & \mathcal{I}(\pi) := \int c\pi. \end{array} \tag{5.4.1}$$

The *Monge-Kantorovich minimization problem* is to find the minimum of $\mathcal{I}$ over all transference plans.

If we consider a μ-measurable *transport map* $T : X \to Y$ such that $T_\sharp \mu = \nu$, the functional (5.4.1) becomes

$$\mathcal{I}(T) := \mathcal{I}\big((Id \times T)_\sharp \mu\big) = \int c(x, T(x))\mu(dx).$$

The minimum problem over all T is called *Monge minimization problem.*

The Kantorovich problem admits a (pre) dual formulation.

Definition 5.8. A map $\varphi : X \to \mathbb{R} \cup \{-\infty\}$ is said to be *c-concave* if it is not identically $-\infty$ and there exists $\psi : Y \to \mathbb{R} \cup \{-\infty\}$, $\psi \not\equiv -\infty$, such that

$$\varphi(x) = \inf_{y \in Y} \big\{c(x,y) - \psi(y)\big\}.$$

The *c-transform* of φ is the function

$$\varphi^c(y) := \inf_{x \in X} \{c(x,y) - \varphi(x)\}. \tag{5.4.2}$$

The *c-superdifferential* $\partial^c \varphi$ of φ is the subset of $X \times Y$ defined by

$$\partial^c \varphi := \Big\{(x,y) : c(x,y) - \varphi(x) \le c(z,y) - \varphi(z)\ \forall z \in X\Big\} \subset X \times Y. \tag{5.4.3}$$

Definition 5.9. A set $\Gamma \subset X \times Y$ is said to be *c-cyclically monotone* if, for any $n \in \mathbb{N}$ and for any family $(x_0, y_0), \dots, (x_n, y_n)$ of points of Γ, the following inequality holds:

$$\sum_{i=0}^n c(x_i, y_i) \le \sum_{i=0}^n c(x_{i+1}, y_i),$$

where $x_{n+1} = x_0$.

A transference plan is said to be *c-cyclically monotone* if it is concentrated on a c-cyclically monotone set.

Consider the set

$$\Phi_c := \Big\{(\varphi, \psi) \in L^1(\mu) \times L^1(\nu) : \varphi(x) + \psi(y) \le c(x,y)\Big\}. \tag{5.4.4}$$

Define for all $(\varphi, \psi) \in \Phi_c$ the functional

$$J(\varphi, \psi) := \int \varphi\mu + \int \psi\nu. \tag{5.4.5}$$

The following is a well known result (see Theorem 5.10 of [31]).

Theorem 5.10 (Kantorovich Duality)**.** *Let X and Y be Polish spaces, let $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$, and let $c : X \times Y \to [0, +\infty]$ be lower semicontinuous. Then the following holds:*

1. *Kantorovich duality:*
$$\inf_{\pi \in \Pi(\mu,\nu)} \mathcal{I}(\pi) = \sup_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi).$$
Moreover, the infimum on the left-hand side is attained and the right-hand side is also equal to
$$\sup_{(\varphi,\psi)\in\Phi_c\cap C_b} J(\varphi,\psi),$$
where $C_b = C_b(X,\mathbb{R}) \times C_b(Y,\mathbb{R})$.

2. *If* c *is real valued and the optimal cost is finite, then there is a measurable* c*-cyclically monotone set* $\Gamma \subset X \times Y$*, closed if* c *is continuous, such that for any* $\pi \in \Pi(\mu,\nu)$ *the following statements are equivalent:*
   (a) π *is optimal;*
   (b) π *is* c*-cyclically monotone;*
   (c) π *is concentrated on* Γ*;*
   (d) *there exists a* c*-concave function* φ *such that* π*-a.s.* $\varphi(x) + \varphi^c(y) = c(x,y)$.

3. *If moreover*
$$c(x,y) \le c_X(x) + c_Y(y), \qquad c_X \ \mu\text{-integrable},\ c_Y \ \nu\text{-integrable},$$
then the supremum is attained:
$$\sup_{\Phi_c} J = J(\varphi,\varphi^c) = \inf_{\pi\in\Pi(\mu,\nu)} \mathcal{I}(\pi).$$

We recall also that if $-c$ is Souslin, then every optimal transference plan π is concentrated on a c-cyclically monotone set [8].

Bibliography

[1] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems.* Oxford Mathematical Monographs, 2000.

[2] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Birkhäuser Verlag, 2005.

[3] L. Ambrosio and B. Kirchheim. Currents in metric spaces. *Acta Mathematica*, 185(1):1–80, 2000.

[4] L. Ambrosio, B. Kirchheim, and A. Pratelli. Existence of optimal transport maps for crystalline norms. *Duke Math. J.*, 125(2):207–241, 2004.

[5] L. Ambrosio and A. Pratelli. Existence and stability results in the L^1 theory of optimal transportation. In Springer, editor, *Optimal Transportation and Applications*, volume 1813 of *Lecture notes in mathematics*, 2001.

[6] L. Ambrosio and P.Tilli. *Topics on analysis in metric spaces.* Oxford Lecture series in Mathematics and its application, 2004.

[7] K. Bacher and K.T. Sturm. Localization and tensorization properties of the Curvature-Dimension condition for metric measure spaces. *J. Funct. Anal.*, 259(1):28–56, 2010.

[8] S. Bianchini and L. Caravenna. On the extremality, uniqueness and optimality of transference plans. *Bull. Inst. Math. Acad. Sin.(N.S.)*, 4(4):353–454, 2009.

[9] S. Bianchini and F. Cavalletti. The Monge problem for distance cost in geodesic spaces. preprint SISSA 50/2009/M, 2010.

[10] S. Bianchini and M. Gloyer. The Euler-Lagrange equation for a singular variational problem. *Math. Z.*, 2009.

[11] V. I. Bogachev. *Gaussian Measures.* Mathematical surveys and monographs. AMS, 1998.

[12] D. Burago, Y. Burago, and S. Ivanov. *A course in metric geometry.* Graduate studies in mathematics. AMS, 2001.

[13] L. Caffarelli, M. Feldman, and R.J. McCann. Constructing optimal maps for Monge's transport problem as a limit of strictly convex costs. *J. Amer. Math. Soc.*, 15:1–26, 2002.

[14] L. Caravenna. An existence result for the Monge problem in $\mathbb{R}^n$ with norm cost function. *preprint*, 2009.

[15] F. Cavalletti. A strategy for non-strictly convex distance transport cost and the obstacle problem. preprint SISSA 50/2010/M, 2010.

[16] F. Cavalletti. The Monge problem in Wiener space. preprint SISSA 09/2011/M, 2011.

[17] F. Cavalletti and K-T. Sturm. Local $\mathsf{CD}(k,n)$ implies $\mathsf{MCP}(k,n)$. forthcoming, 2011.

[18] T. Champion and L. De Pascale. The Monge problem in $\mathbb{R}^d$. *preprint*, 2009.

[19] L.C. Evans and W. Gangbo. Differential equations methods for the Monge-Kantorovich mass transfer problem. *Current Developments in Mathematics*, pages 65–126, 1997.

[20] M. Feldman and R. McCann. Monge's transport problem on a Riemannian manifold. *Trans. Amer. Math. Soc.*, 354:1667–1697, 2002.

[21] D. H. Fremlin. *Measure Theory*, volume 4. Torres Fremlin, 2002.

[22] D.G. Larman. A compact set of disjoint line segments in $\mathbb{R}^3$ whose end set has positive measure. *Mathematika*, 18:112–125, 1971.

[23] J. Lott and C. Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math.*, 169(3):903–991, 2009.

[24] S.-I. Ohta. On the measure contraction property of metric measure spaces. *Comment. Math. Helv.*, 82:805–828, 2007.

[25] A. M. Srivastava. *A course on Borel sets.* Springer, 1998.

[26] K.T. Sturm. On the geometry of metric measure spaces.I. *Acta Math.*, 196(1):65–131, 2006.

[27] K.T. Sturm. On the geometry of metric measure spaces.II. *Acta Math.*, 196(1):133–177, 2006.

[28] V.N. Sudakov. Geometric problems in the theory of dimensional distributions. *Proc. Steklov Inst. Math.*, 141:1–178, 1979.

[29] N. Trudinger and X.J. Wang. On the Monge mass transfer problem. *Calc. Var. PDE*, 13:19–31, 2001.

[30] C. Villani. *Topics in optimal transportation.* Graduate studies in mathematics. AMS, 2003.

[31] C. Villani. *Optimal transport, old and new.* Springer, 2008.